TI 2018-025/IV
Tinbergen Institute Discussion Paper

Forecasting with Bayesian Vector Autoregressions with Time Variation in the Mean

Marta Banbura[1]
Andries van Vlodrop[2]

1: European Central Bank, Germany
2: VU Amsterdam, the Netherlands

Forecasting with Bayesian Vector Autoregressions with Time Variation in the Mean

Marta Bańbura* Andries van Vlodrop†

March 8, 2018

Abstract

We develop a vector autoregressive model with time variation in the mean and the variance. The unobserved time-varying mean is assumed to follow a random walk and we also link it to long-term Consensus forecasts, similar in spirit to so called democratic priors. The changes in variance are modelled via stochastic volatility. The proposed Gibbs sampler allows the researcher to use a large cross-sectional dimension in a feasible amount of computational time. The slowly changing mean can account for a number of secular developments such as changing inflation expectations, slowing productivity growth or demographics. We show the good forecasting performance of the model relative to popular alternatives, including standard Bayesian VARs with Minnesota priors, VARs with democratic priors and standard time-varying parameter VARs for the euro area, the United States and Japan. In particular, incorporating survey forecast information helps to reduce the uncertainty about the unconditional mean and along with the time variation improves the long-run forecasting performance of the VAR models.

Keywords: Consensus forecasts, forecast evaluation, large cross-sections, state space models.

JEL Classification: C11, C32, C53, C55, E37.

* European Central Bank, e-mail: marta.banbura@ecb.europa.eu.

† Vrije Universiteit Amsterdam, e-mail: a.c.van.vlodrop@vu.nl.

We would like to thank Francisco Blasques, Adam Elbourne, Siem Jan Koopman, André Lucas, Emanuel Mönch, Inske Pirschel, Giorgio Primiceri and the participants at Computing in Economics and Finance 2017, the Bundesbank Workshop on Forecasting 2017, the 5th INFER Workshop on Applied Macroeconometrics, the 18th IWH-CIREQ-GW Macroeconometric Workshop, Computational and Financial Econometrics 2017, a CPB seminar and an ECB seminar, for their comments.

The opinions in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank and the Eurosystem.

1 Introduction

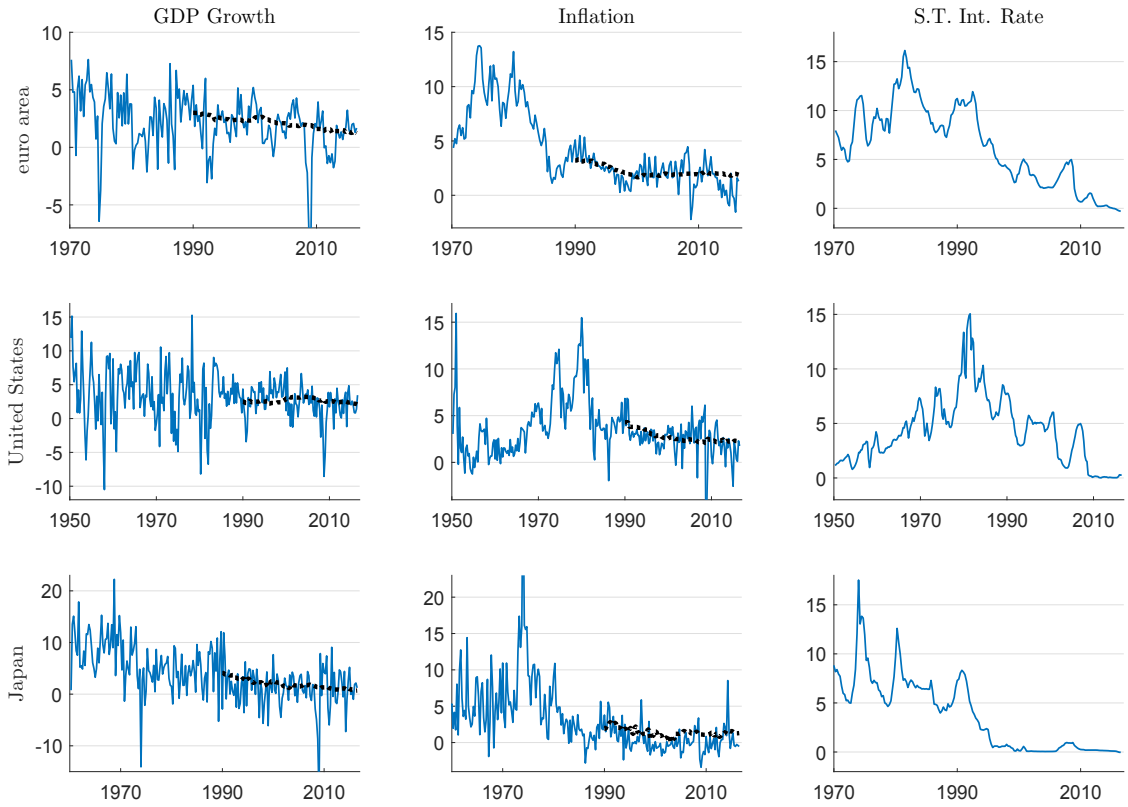
We develop and evaluate an extension of vector autoregressive (VAR) models that allows for variation in the level of the unconditional mean over time. In a standard VAR representation a time-invariant filter is applied to variables in deviation from their *time-invariant* unconditional mean. Our approach relaxes the latter assumption by considering a constant coefficient VAR for variables in deviation from a *local* mean vector, which we allow to vary over time as a random walk. We also allow additional "signal" on the unobserved local mean by linking it to some observable variable, in our case long-term Consensus forecasts, via a measurement equation. Finally, we allow for time variation in the variances by assuming stochastic volatility for all shocks.

The motivation for the time variation in the mean originates from relevant "low-frequency" changes that can be observed in a number of key macroeconomic variables. For example, over the last decades many advanced economies, including the euro area, the United States and Japan, experienced declining inflation and interest rates, reflecting, among other things, structural changes to the conduct of monetary policy and to the formation of inflation expectations. Also average growth rates of output have been changing over time in many countries, related inter alia to developments in demographics or productivity. Figure 1 plots the developments of these key macroeconomic variables together with corresponding long-term forecasts from Consensus Economics. First, the time series shown in the figure seem to invalidate approaches which take the unconditional mean of these series to be constant over time. Second, the long-term forecasts appear to track well low-frequency movements in GDP growth and inflation.

The common approach in the literature to handle structural changes in the VAR context is to allow all the VAR coefficients to vary over time (as random walks). Examples of such time-varying parameter VARs (TVP-VARs) include Cogley and Sargent (2005) and Primiceri (2005). Such general case of time variation implies, however, a large number of time-varying parameters to perform inference on. This becomes problematic as the number of variables included in the VAR increases. Computational considerations arise from the time requirements of the MCMC sampler.[1] Also, heavy parameterisation might lead to overfitting and large uncertainty associated with the parameter estimates. Typically, TVP-VARs as specified in Cogley and Sargent (2001, 2005) and Primiceri (2005) only include three variables and two autoregressive lags. Several approaches have been proposed to address these problems. For example, Koop and Korobilis (2013) propose an approximation to the posterior distribution by using forgetting factors to reduce the computational burden related to repeated Kalman filter runs. Eisenstat, Chan, and Strachan (2016) propose stochastic model specification search for TVP-VARs with stochastic volatility in order to deal with overfitting and the typically imprecise inference in these highly parameterised models (see also Frühwirth-Schnatter and Wagner, 2010; Belmonte, Koop, and Korobilis, 2014). de Wind and Gambetti (2014) introduce cross-equation restrictions on the time variation.

[1] This is because more MCMC draws are needed, but also the MCMC steps itself become more demanding, thereby explosively increasing the required computational time.

Figure 1: Key macroeconomic variables and their long-term survey forecasts



Notes: Blue lines show GDP growth, inflation and the short-term interest rate (in columns) for the euro area, the United States and Japan (in rows). The dotted lines show the long-term survey forecasts for GDP growth and inflation from Consensus Economics.

The approach proposed in this paper allows for time variation in the mean and the variance of the variables but assumes that the coefficients of the VAR are constant over time. Therefore it allows the researcher to consider a larger number of variables than in a "fully" time-varying version and potentially limits estimation uncertainty. Similar approaches were recently adopted in Garnier, Mertens, and Nelson (2015), Crump, Eusepi, and Moench (2016) and Andrle and Bruha (2017). In particular, Garnier, Mertens, and Nelson (2015) fit such a VAR to inflation rates and they associate the local mean with the inflation "trend", in the spirit of the Beveridge-Nelson decomposition since it coincides with the infinite horizon forecast. They find little evidence of time variation in the VAR coefficients for the deviations of inflation rates from the trend.

As an additional element, we link the local mean with the long-term Consensus forecasts via a measurement equation. Such "anchoring" allows for adding off-model information and for reducing the dispersion of the posterior distribution of the local mean. The information value of survey forecasts, such as of those collected by Consensus Economics, for model-based forecasting has been

shown in many studies. Ang, Bekaert, and Wei (2007) and Faust and Wright (2013) show that survey forecasts by themselves already have good forecasting properties for inflation compared to model based forecasts. Wright (2013) shows the improved forecasting performance from including long-term survey forecasts in VAR models via "democratic priors" on the unconditional mean. Clark and McCracken (2006) and Clark and Doh (2014) show that detrending inflation using survey forecasts can lead to better forecasting performance of (V)AR models compared to non-detrended versions. Several studies exploit the information from survey forecasts to estimate inflation expectations or trends (see e.g. Kozicki and Tinsley, 2012; Mertens, 2016; Chan, Clark, and Koop, 2017; Lenza and Jarociński, 2016). Crump, Eusepi, and Moench (2016) use a term structure of professional forecasts for GDP growth, inflation and interest rates to inform the low-frequency evolution of these variables.

Besides time variation in the mean we introduce time variation in the variances, precisely we allow for stochastic volatility. First, we assume stochastic volatility in the VAR innovations. A number of studies provides empirical evidence that adding stochastic volatility to VARs improves in-sample fit and forecast accuracy (see e.g. Clark, 2011), also in the case of large VARs (Carriero, Clark, and Marcellino, 2016a,b). Moreover, Sims (2001) and Stock (2001) argue that when adding stochastic volatility outliers are down-weighted and a part of the time variation in the coefficients of TVP-VARs might disappear. Secondly, we assume stochastic volatility for the innovations in the local mean process. As a result, the model nests the univariate unobserved components with stochastic volatility (UCSV) model of Stock and Watson (2007), who document the importance of allowing for stochastic volatility in trend inflation. The UCSV model has been shown to have good forecasting properties for inflation and has become a popular benchmark in forecast evaluations. Finally, we also consider stochastic volatility in the equation relating the local mean to long-term survey forecasts, which can potentially account for periods in which the local mean differs more strongly from the survey forecasts.

We adopt the Bayesian approach for inference and we approximate the posterior distribution of the parameters using a Gibbs sampler. Conveniently, the model allows a linear Gaussian state space representation.[2] The Gibbs sampling step for inference on the local mean process is then performed using a version of the simulation smoother of Carter and Kohn (1994). A key feature of the local mean approach is that it admits the use of a large cross-sectional dimension.[3] For the coefficients of the VAR we take the Minnesota-type priors of Doan, Litterman, and Sims (1984). Further, we adopt the triangularisation method of Carriero, Clark, and Marcellino (2016b), which allows estimating VARs with a large cross-sectional dimension also when stochastic volatility is assumed for the VAR innovations.

We analyse the in-sample properties of the model and evaluate it in an out-of-sample forecasting

[2]To be more precise, this holds when we condition on the stochastic volatility processes.

[3]Previous studies indicate that VARs with a larger cross-sectional dimension might be preferred over smaller VAR models if an appropriate degree of coefficient shrinkage is adopted (see e.g. Bańbura, Giannone, and Reichlin, 2010).

exercise for forecast horizons ranging from one-quarter to three years. We focus on the results for the euro area but we also check the robustness of the results and the scalability of the model using datasets for the United States and Japan. Given the availability of Consensus forecasts, the evaluation sample runs from 2000Q1 until 2016Q3. The benchmark models include standard VARs with Minnesota-type priors, VARs with democratic priors, standard TVP-VARs and the UCSV model. The forecast evaluations indicate favourable forecasting performance of the model compared to the benchmarks, in particular at longer horizons. In more detail, it appears that the improved forecasting performance of the model is due *both* to the inclusion of the long-term survey forecast information (as the "democratic prior" model is also competitive and a version of the local mean VAR without the link to survey forecasts performs worse) *and* to allowing for time variation in the mean (as the model often improves over the "democratic prior" model). The good forecasting performance relative to TVP-VARs could indicate that the model captures better the relevant dimension of time variation, is subject to reduced estimation uncertainty via a lighter parameterisation *and/or* has an advantage from the reliance on off-model information. Importantly, survey information reduces the dispersion of the posterior distribution of the local (unconditional) mean. This results in smaller forecast revisions between consecutive forecast updates, which is an important consideration for practitioners. The off-model information also seems to be key when enlarging the cross-sectional dimension of the local mean VAR model.

This paper contributes to a growing literature on how best to model time variation in macro time series models in a forecasting context. Recent studies include D'Agostino, Gambetti, and Giannone (2013) who report better forecasting performance of TVP-VARs over constant parameter versions for inflation (but not for unemployment and the interest rate). Barnett, Mumtaz, and Theodoridis (2014) show that some models with time variation tend to outperform constant parameter models but not consistently across variables and evaluation samples. Aastveit, Carriero, Clark, and Marcellino (2017) document evidence for time variation in the VAR framework and evaluate a range of constant parameter VARs and VARs allowing for some time variation. They show that no model consistently beats the others; TVP-VARs tend to work well with smaller systems but constant parameter models are in general hard to beat. Dijk, Koopman, Wel, and Wright (2014) stress the importance of allowing for shifting endpoints in yield curve factors when forecasting interest rates. In addition, they argue that "anchoring" such endpoints by exploiting the information from survey forecasts yields good forecasting performance. Stock and Watson (2010) recommend a Phillips curve for inflation in deviation from a stochastic trend and find that the reaction to a gap measure is stable for such deviation. Stock and Watson (2012) indicate the importance of first detrending the data before fitting a (constant coefficient) dynamic factor model (DFM). Also Antolin-Diaz, Drechsel, and Petrella (2017) show that allowing for a time-varying intercept for GDP within a DFM leads to improved forecasting performance. Götz and Hauzenberger (2017) allow for time-varying intercepts in a mixed frequency VAR and they show advantages with respect to both constant-parameter and fully time-varying versions.

The rest of this paper is organised as follows. Section 2 presents the local mean VAR model we study in this paper. In Section 3 we describe the data, provide implementation details and

present some in-sample estimation results. Section 4 contains the description and the results of the forecast evaluations. Section 5 provides a more detailed analysis of the forecasts for the euro area, furthermore this section reports impulse response functions for different models and discusses forecasting conditional on scenarios for long-term expectations. Section 6 contains some sensitivity analysis. A conclusion follows in Section 7.

2 Model description

2.1 Local mean VAR model

The local mean VAR model is a generalisation of the standard VAR model with a time-invariant unconditional mean. Let y_t denote the M-dimensional vector containing the variables of interest at time t. Further, the vector ψ_t denotes the *local mean* of these variables at time t. Consider the VAR for the variables in deviation from their local mean:

$$y_t - \psi_t = \sum_{k=1}^{p} B_k \left(y_{t-k} - \psi_{t-k}\right) + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, H_t\right), \tag{1}$$

where H_t represents the (time-varying) covariance matrix of the innovations. We assume that the VAR coefficients $B_1, \ldots, B_p$ are time-*invariant* and that they satisfy the usual stability condition required for stationary VARs.

We model the process for the local mean vector, ψ_t, as univariate random walk processes for each of its elements:

$$\psi_t = \psi_{t-1} + \eta_t, \quad \eta_t \sim N(0, V_t), \tag{2}$$

where V_t is diagonal: $V_t = \text{diag}(\sigma^2_{V,1,t}, \ldots, \sigma^2_{V,M,t})$. Under these assumptions we have that:

$$\lim_{h\to\infty} E_t\left[y_{t+h} \mid \psi_t\right] = \psi_t, \tag{3}$$

where E_t denotes the expectation conditional on the information set available at time t. The interpretation of (3) is that the point forecasts from the local mean VAR model in (1) and (2) converge, as the forecast horizon increases, towards the local mean at time t.

2.2 Including information from long-term survey forecasts

Let z_t denote the M_z-dimensional vector of long-term survey forecasts at time t for $t = t_z, \ldots, T$. We note that the survey forecasts might be unavailable at the beginning of the sample, therefore we indicate by t_z the first point in time when these are available.[4] We exploit the information

[4]For notational simplicity we use here a single t_z for all the survey forecasts. The modification to the case with individual survey forecasts starting in different points in time is straightforward.

from survey forecasts on the long-term developments in the variables of interest by linking the unobserved local mean to such forecasts through a measurement equation:

$$z_t = \psi_t + g_t, \quad g_t \sim N\left(0, G_t\right), \tag{4}$$

where G_t is diagonal: $G_t = \text{diag}(\sigma^2_{G,1,t}, \ldots, \sigma^2_{G,M_z,t})$, since we assume no cross-sectional dependence between the measurement equations defined in (4).

Survey forecasts might only be available for a selection of the variables, $M_z \leq M$. In case of such partial coverage we use a selection matrix, P_ψ, that selects the M_z elements in the ψ_t vector corresponding to the variables for which the survey information is available and we replace ψ_t by $P_\psi \psi_t$ in (4), see Section 3.2 for an example.

Compared to the detrending of VAR variables using survey forecasts, one of the advantages of the approach proposed here is that shorter samples and other cases of missing observations in the survey forecasts can be easily handled using Kalman filtering techniques.

We shall refer to the local mean VAR model augmented with survey forecasts as the *survey local mean (SLM)* VAR model. For brevity we shall often omit the VAR part and refer to it as the SLM model. We note that the SLM model described by equations (1)-(2) and (4) can conveniently be cast in a state space form (see Appendix A for the details).

2.3 Stochastic volatility

Stochastic volatility is assumed for the innovations in each of the equations (1)-(2) and (4). For H_t in (1) we follow Carriero, Clark, and Marcellino (2016b) and decompose it as:

$$H_t = A^{-1} \Lambda_t \left(A^{-1}\right)', \quad \Lambda_t = \text{diag}\left(\sigma^2_{H,1,t}, \ldots, \sigma^2_{H,M,t}\right), \tag{5}$$

where A is a (time-invariant) lower triangular matrix with ones on its diagonal and Λ_t is a diagonal matrix with the time-varying variances of the shocks on the diagonal. The covariance matrices V_t and G_t from (2) and (4) are both taken to be diagonal. What remains is to specify the stochastic processes for the variances. We follow the standard approach and take logarithmic transformations and model these as random walk processes:[5]

$$\ln \sigma^2_{H,i,t} = \ln \sigma^2_{H,i,t-1} + u_{H,i,t}, \quad u_{H,i,t} \sim N(0, \phi_{H,i}), \quad i = 1, \ldots, M, \tag{6}$$

$$\ln \sigma^2_{V,i,t} = \ln \sigma^2_{V,i,t-1} + u_{V,i,t}, \quad u_{V,i,t} \sim N(0, \phi_{V,i}), \quad i = 1, \ldots, M, \tag{7}$$

$$\ln \sigma^2_{G,i,t} = \ln \sigma^2_{G,i,t-1} + u_{G,i,t}, \quad u_{G,i,t} \sim N(0, \phi_{G,i}), \quad i = 1, \ldots, M_z. \tag{8}$$

[5]The random walk specification is more parsimonious compared to a mean-reverting version and therefore might be preferable for forecasting. Clark and Ravazzolo (2015) evaluate different forms of time-varying volatility in the Bayesian VAR framework and find that stochastic volatility dominates other specifications in terms of forecasting performance and that there is little difference between the mean-reverting and random walk specifications.

2.4 Prior distributions

The set of parameters in the model is $\{\{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T\}$, where the superscript T denotes the full time series of a parameter and $B = [B_1 : \ldots : B_p]'$ collects the VAR coefficients. We base our specifications of the prior distributions for these parameters mainly on common conventions in the literature on Bayesian VARs.

For the VAR coefficients we take an independent normal prior:[6]

$$\text{vec}\,(B) \sim N\left(\underline{\mu}_B, \underline{\Omega}_B\right), \tag{9}$$

where we specify the prior variances in the diagonal matrix $\underline{\Omega}_B$ following the ideas behind the Minnesota prior of Doan, Litterman, and Sims (1984). Precisely, we impose more shrinkage on more distant lags and we make the distinction between the coefficients relating variables to their own lags (the diagonal elements of the B_k matrices) and the coefficients relating to lags of other variables (the off-diagonal elements). For the diagonal coefficients we set the prior variances as:

$$\text{var}\left(B_l^{(i,i)}\right) = \frac{\lambda^2}{l^2}, \quad i = 1, \ldots, M, \quad l = 1, \ldots, p, \tag{10}$$

and for the variances of the off-diagonal VAR coefficients we take:

$$\text{var}\left(B_l^{(i,j)}\right) = \tau \frac{\lambda^2}{l^2} \frac{\sigma_i^2}{\sigma_j^2}, \quad i, j = 1, \ldots, M, \quad i \neq j, \quad l = 1, \ldots, p. \tag{11}$$

The hyperparameter λ controls the degree of overall shrinkage towards the prior means, i.e. how important the priors are compared to the information in the data. For $\lambda \to \infty$ the priors become uninformative, while as $\lambda \to 0$ the posterior distributions of the VAR coefficients are increasingly shrunk towards the prior means. With the hyperparameter τ more shrinkage is imposed on the off-diagonal elements. Furthermore, the prior on the off-diagonal elements takes into account the relative scales of the variables through σ_i^2/σ_j^2.

For the local mean process $\psi_{p+1}, \ldots, \psi_T$ we set a Gaussian prior on the initial value ψ_{p+1}:

$$\psi_{p+1} \sim N\left(\underline{\mu}_\psi, \underline{\Omega}_\psi\right). \tag{12}$$

Turning to the prior distributions for the parameters related to the covariance matrices in the VAR, for the A matrix in the decomposition in (5) we specify the priors row by row as in Cogley and Sargent (2005). Let a_i, $i = 2, \ldots, M$, denote the i^{th} row of the below diagonal part of the matrix A. The vector a_i is of dimension $i - 1$ and we use the following prior distributions:

$$a_i \sim N\left(\underline{\mu}_{a,i}, \underline{\Omega}_{a,i}\right), \quad i = 2, \ldots, M. \tag{13}$$

[6] We refer to Kadiyala and Karlsson (1997) and Karlsson (2013) for a discussion of priors on B.

For the stochastic volatility processes we specify the distribution of the initial value and the prior distributions for the variance parameters ϕ governing the amount of time variation in the process as follows:

$$\ln \sigma^2_{H,i,p+1} \sim \mathrm{N}\left(0, v_H\right), \quad \phi_{H,i} \sim \mathrm{IG}\left(d_{\phi_H} \times \underline{\phi}_{H,i}, d_{\phi_H}\right), \; i = 1, \ldots, M, \tag{14}$$

$$\ln \sigma^2_{V,i,p+2} \sim \mathrm{N}\left(0, v_V\right), \quad \phi_{V,i} \sim \mathrm{IG}\left(d_{\phi_V} \times \underline{\phi}_{V,i}, d_{\phi_V}\right), \; i = 1, \ldots, M, \tag{15}$$

$$\ln \sigma^2_{G,i,t_z} \sim \mathrm{N}\left(0, v_G\right), \quad \phi_{G,i} \sim \mathrm{IG}\left(d_{\phi_G} \times \underline{\phi}_{G,i}, d_{\phi_G}\right), \; i = 1, \ldots, M_z, \tag{16}$$

2.5 Gibbs sampler

We specify a Gibbs sampling algorithm to approximate the posterior distribution of the parameters, the details of which are provided in Appendix A. We emphasise two elements of our inference approach. First, it is computationally feasible for a large cross-section of time series. Notably, we implement the recently proposed triangularisation method of Carriero, Clark, and Marcellino (2016b) to sample the VAR coefficients B equation by equation, which allows for having both a large cross-sectional dimension as well as stochastic volatility in the VAR innovations. Second, the inference on the local mean process can be performed through standard simulation smoothing methods, where we use a version of the simulation smoother of Carter and Kohn (1994).

3 Data, implementation details and in-sample results

3.1 Data

Our dataset consists of quarterly observations for key macroeconomic variables as well as survey based long-term forecasts for some of those variables for the euro area, the United States and Japan. For each economy we consider real GDP, investment and consumption growth, as well as inflation and the nominal short-term interest rate, for the period 1970Q2-2016Q3.[7] Detailed information on the dataset is provided in Table A1 in Appendix B. Most of the results reported in this paper are for a three-variable system with real GDP growth, inflation and the short-term interest rate for the euro area.

The long-term survey forecasts are provided by Consensus Economics and correspond to the average 6 to 10 year ahead expectations. Only a very short time series is available for the short-term interest rate forecasts for the euro area and we do not use it in the analysis. Prior to 2003 the survey forecasts for the euro area are derived by aggregating the forecasts for the (available) euro area countries. The forecasts are released in April and October of each year. We assign the April value to Q1 and Q2 and the October value to Q3 and Q4.[8]

3.2 Implementation details

The applications in this paper involve quarterly data and we make the common choice to include $p = 4$ lags in the VAR. For the hyperparameters λ and τ in the Minnesota prior on the VAR coefficients B we take the usual values in the literature, i.e. $\lambda = 0.2$ and $\tau = 0.5$.[9] Regarding the variance scaling parameters σ_i^2, we take them as the residual variances of univariate AR(4) models estimated on a training sample. For the prior mean for B we take $\underline{\mu}_B = 0$ as is commonly done for stationary VARs. We further impose stationarity of the VAR equation in (1) by rejecting instable draws of B.

For the local mean process we set the prior mean $\underline{\mu}_\psi$ for ψ_{p+1} equal to the average values of the variables in the training sample. We then obtain a diffuse initialisation by setting $\underline{\Omega}_\psi = 1,000 \times I_M$.

For the stochastic volatility processes we specify diffuse initial conditions. For the hyperparameters in the inverse Gamma prior on the variance of the innovations we take $d_\phi = 10$ and $\underline{\phi} = 0.01$

[7] All the data considered have the same time stamp, therefore we do not take into account data revisions over time.

[8] Alternatively we could treat the data as missing in two quarters of each year. Forecast accuracy is comparable when following this approach. The results are available upon request. As of 2014 long-term Consensus forecasts are available on a quarterly basis.

[9] For the larger cross-section of 15 variables we increase the shrinkage as advocated by e.g. Bańbura, Giannone, and Reichlin (2010), i.e. we set $\lambda = 0.1$.

for each equation. For the rows of the A matrix we specify an uninformative prior, i.e. $\underline{\mu}_{a,i} = \underline{0}$ and $\underline{\Omega}_{a,i} = 1,000 \times I_{i-1}$ for $i = 2, \ldots, M$.

For the Cholesky decomposition of the covariance matrix H_t in (5) we need to decide on the ordering of the variables, which matters slightly for the posterior distribution, as mentioned by e.g. Primiceri (2005). Our choice corresponds to the common practice in the literature on Bayesian VARs with stochastic volatility, namely we follow the insight of slow and fast-moving variables of Bernanke, Boivin, and Eliasz (2005). The ordering is as displayed in Table A1.

We note that the survey forecast data is not available for the early period, where we treat them as missing data (see Appendix A for details). Further, as mentioned above, our data set does not include long-term survey forecasts of the short-term interest rate. We specify the selection matrix P_ψ such that all the local means, except those for the short-term interest rate, are linked to long-term survey forecasts. For example, the equation relating survey forecasts to the local mean in (4) for the three-variable system is as follows:

$$z_t = P_\psi \psi_t + g_t, \quad P_\psi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{17}$$

We start the Gibbs sampling algorithm with a burn-in sample of $10,000$ draws which we then discard and consider the following $25,000$ draws for inference purposes.[10] To generate forecasts, for each Gibbs draw (generated over the estimation sample) we produce forecast paths in two steps: in the first step we simulate the time-varying parameters out-of-sample; in the second step the paths of the variables (innovations) are simulated. To increase the accuracy, we repeat these two steps 25 times for each of the Gibbs draws.

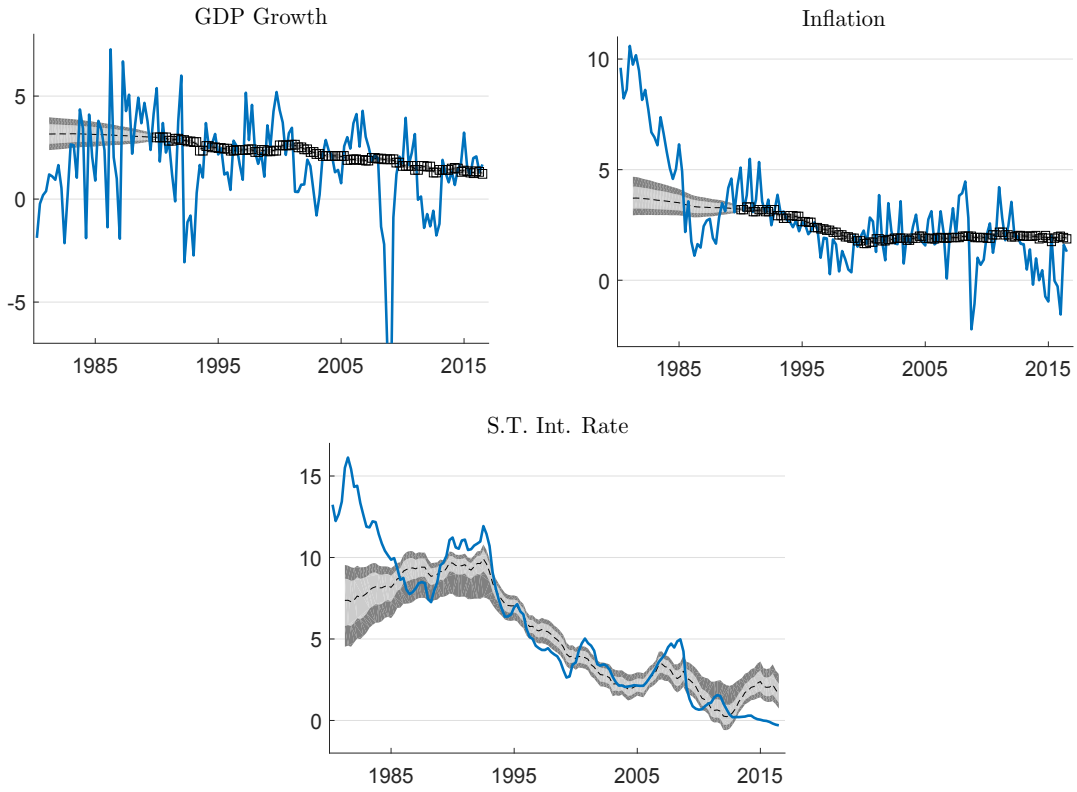
3.3 In-sample results

In this section we look at the posterior distributions of the local mean and the stochastic volatility processes in the SLM model for the three-variable system (real GDP growth, inflation and the short-term interest rate, in this order) for the euro area. The training sample runs from 1970Q2 until 1981Q1 and the estimation period (for which we show the in-sample posterior distribution) is 1980Q2-2016Q3. As explained above, only the local means for GDP growth and inflation are anchored by the long-term survey forecasts.

Figure 2 depicts the posterior distribution of the local mean process. The posterior median and the 68% and 90% (symmetric) coverage bands are shown together with the data and the survey

[10] Casual checks confirm that a burn-in sample of $10,000$ draws is enough to reach convergence. Furthermore, the inefficiency factors show satisfactory mixing properties of the proposed Gibbs sampling algorithm for the three-variable VARs, but also for the large cross-section of 15 variables. This indicates that the Gibbs sampling algorithm is able to handle also larger cross-sections efficiently. The results are available upon request.

Figure 2: Posterior distribution for the local mean in the SLM model

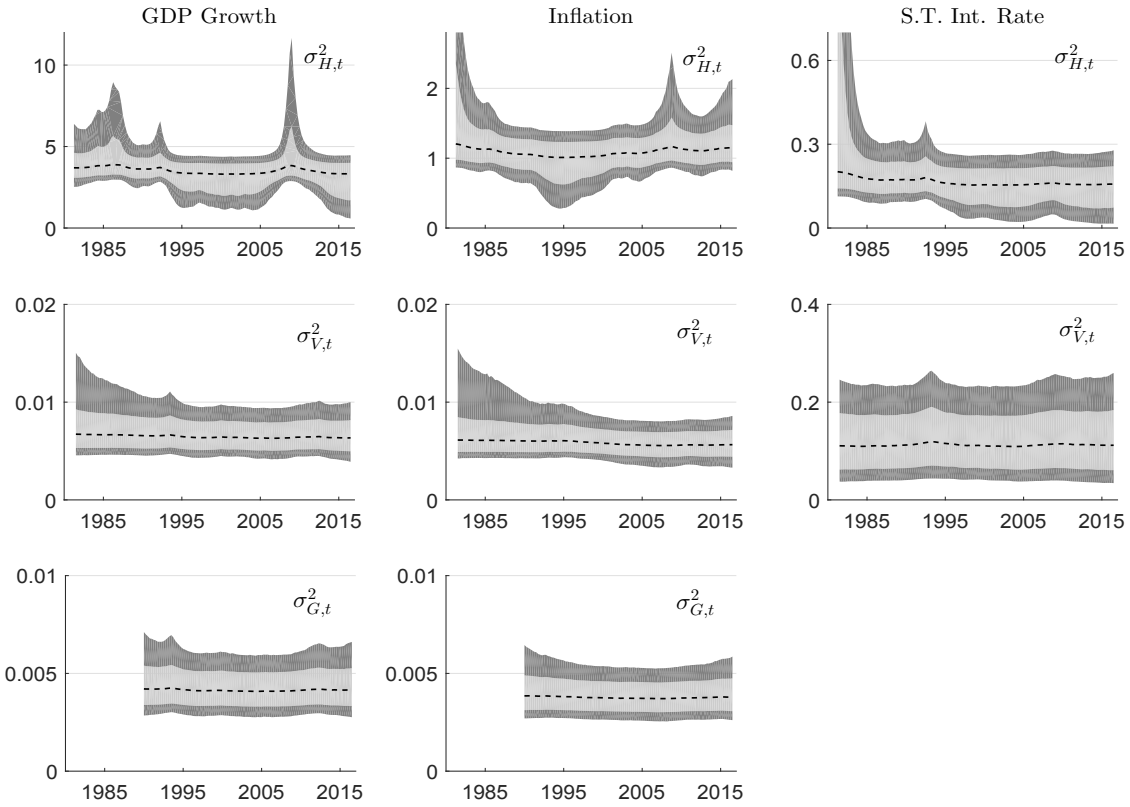


Notes: In-sample estimates of the local mean process from the SLM model with three variables. The posterior distribution is summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. The blue line corresponds to the data and the squares to the long-term survey forecasts.

forecasts. For GDP growth and inflation there is a strong anchoring of the local mean process towards the survey forecasts when the latter become available.[11] Also, over the full sample period the local mean process captures well a slowly declining trend in these variables. For the short-term interest rate, as there is no anchoring by long-term forecasts, the posterior distribution of the local mean process is wider and there is more pronounced time variation (the local mean "tracks" the data more). It is noteworthy that at the end of the sample the local mean process moves upwards, while the short-term interest rate remains around zero. This appears to be the result of the fact that the actual inflation has remained below its local mean for some time and that the estimated time-invariant VAR relationships suggest that this should be associated with the interest rate likewise remaining below its corresponding local mean (consistent with a Taylor rule principle).

[11]Note that we have selected rather diffuse priors on the variances in (1) and (4), therefore the strong anchoring is the result of the inference in the model (rather than being imposed).

Figure 3: Posterior distribution for the stochastic volatility in the SLM model



Notes: In-sample estimates of the stochastic volatility processes. The rows report the posterior distributions of the stochastic volatility for respectively: the shocks in the VAR; the error in the local mean process; and the error in the measurement equation linking the local mean to the survey forecasts. The posterior distribution is summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands.

Figure 3 shows the posterior distributions of the stochastic volatility processes. The top panels summarise the posterior for the variances of the shocks in the VAR in (1). Some time variation can be found in the shock variances for GDP growth and inflation. For example, during the recent financial crisis of 2008 these variances increased strongly. For the short-term interest rate there does not seem to be much time variation in the variance.[12] For the posterior distributions of the innovation variances in the random walk equation for the local mean in (2) and in the second measurement equation given by (4) we find less evidence of time variation, especially in the second part of the sample. However, to keep the model specification general we allow for stochastic volatility in the errors in each of relationships (1)-(2) and (4).

[12]The setting for the interest rate is also different. The local mean is not anchored and can move more freely, thereby potentially taking away a part of the time variation in the variance. In addition, as the interest rate is ordered as last, it is also affected by the other two shocks (due to the Cholesky decomposition). Finally, the interest rate is characterised by less volatility than GDP growth and inflation.

4 Forecast evaluation

4.1 The forecasting exercise

The out-of-sample period on which the forecasts are evaluated is 2000Q1-2016Q3. We use the first 10-year period, 1970Q2-1981Q1, as a training sample. For the first forecast we use data from 1980Q2-1999Q4 to run the Gibbs sampler, we then recursively expand this estimation window by one quarter for each new forecast in the evaluation sample.[13] Comparisons based on rolling window estimation are considered in the sensitivity analysis in Section 6.

We consider forecast horizons of 1, 4, 8 and 12 quarters. Given the convergence of the point forecasts to the unconditional or local mean, the relative accuracy over longer forecast horizons is of particular interest for evaluating the SLM model.

Forecast accuracy is evaluated using two measures. For the point forecasts we compute the root mean squared error (RMSE), comparing the mean of the simulated forecasts with the out-of-sample observations. The density forecasts, constructed by fitting a Gaussian kernel to the simulated forecasts, are evaluated with the log predictive score (LPS). Alternative measures of accuracy are considered in Section 6.

The target variables in the forecast evaluations are the annualised GDP growth rate, the annualised inflation rate and the level of the short-term interest rate. To abstract from volatility, for GDP growth and inflation for the horizons of 4, 8 and 12 quarters we take the annualised averages, over respectively the first, second and third year (for the forecast and the target variable). For the short-term interest rate we forecast the level for each horizon.

In order to evaluate whether the SLM model can be applied also for larger datasets we look at three different sizes of the cross-section in the VARs. The smallest dataset contains the three target variables. To obtain a medium-size dataset we add investment and consumption, resulting in a five-variable VAR. We construct the largest dataset by merging the medium-size datasets for the three economies, obtaining a cross-sectional dimension of 15 variables.[14] For a VAR with time-varying parameters and stochastic volatility, 15 variables is generally considered to be a relatively large dataset.[15]

In this section we consider pure out-of-sample forecasting with the SLM model. We defer to Section 5.5 a discussion on how the SLM model can be used for forecasting conditional on a

[13] We use a single data vintage and therefore forecast comparisons abstract from the issue of real-time data revisions.

[14] The ordering is as follows: for each variable (category) we put first the series for the United States, followed by the euro area and Japan, we then order the categories as in Table A1.

[15] We note that, at least from a computational point of view, the SLM model is well able to handle larger datasets than those we use here.

future scenario for the long-term survey forecasts.

In terms of time variation the SLM model falls in between the constant parameter VARs and the TVP-VARs. Accordingly, we split the forecasting exercises in two parts. All models allow for time variation in the variance of the shocks via stochastic volatility. The results for implementations with constant variance are reported in Section 6.

4.2 Comparison to constant coefficient VARs

4.2.1 Benchmark models

The first benchmark model is the standard Bayesian VAR with constant coefficients and intercept:

$$y_t = c + \sum_{k=1}^{p} B_k y_{t-k} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, H_t\right), \tag{18}$$

with Minnesota-type priors for c and B_k's. In what follows we label this model as the constant coefficients (CC) VAR model. The SLM model extends on this benchmark by allowing for time variation in the unconditional mean as well as by anchoring it to the long-term survey forecasts.

To better understand the role of the two elements just mentioned we also consider the democratic prior (DP) approach of Wright (2013). First, under the stability conditions on B_k mentioned in Section 2.1 we can reparameterise (18) into:

$$y_t - \mu = \sum_{k=1}^{p} B_k \left(y_{t-k} - \mu\right) + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, H_t\right), \tag{19}$$

where μ represents the time-invariant unconditional mean. Villani (2009) proposes to put an informative prior on μ and the DP approach uses the long-term survey forecasts as the mean of this prior. As a result, the DP benchmark model allows for an anchoring of the unconditional mean to survey forecasts, but assumes the unconditional mean to be time-invariant.[16]

To put the two benchmark models a priori on an equal footing with the SLM model the prior settings for B_k's and for the parameters related to stochastic volatility are the same as in the latter model. The prior for c in specification (18) is diffuse. For the DP approach, when the survey forecast is available it is taken as the prior mean for the corresponding element in μ and the prior variance is set at 0.05. When, instead, no survey forecast is available we take a diffuse prior. This coincides with the standard implementation as in Wright (2013) (with the exception of stochastic volatility).

[16] Furthermore, for a given estimation sample, it only relates the unconditional mean to survey forecasts at the end of the sample, while ignoring the historical values of these forecasts. One could even be more critical and argue that the way survey forecasts enter as a prior in these models is not purely Bayesian, because survey respondents can observe the past values of the variables before stating their forecasts and therefore the prior information comes partly from the estimation sample.

4.2.2 Results

Table 1: Forecast evaluation results for three-variable VARs

		euro area			United States			Japan		
RMSE		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.26	0.96^{***}	$\mathbf{0.94}^{***}$	2.62	0.98	**0.98**	4.43	**1.00**	1.01
Growth	4Q	2.17	0.94^{**}	$\mathbf{0.86}^{*}$	2.03	$\mathbf{0.90}^{**}$	0.92^{*}	2.57	**0.99**	1.00
	8Q	2.47	0.94^{**}	$\mathbf{0.87}^{**}$	2.21	0.92^{**}	$\mathbf{0.89}^{*}$	**2.52**	1.00	1.01
	12Q	2.51	0.94^{***}	$\mathbf{0.90}^{***}$	2.23	0.93^{**}	$\mathbf{0.93}^{**}$	**2.48**	$1.02^{\circ\circ\circ}$	1.01°
Inflation	1Q	1.27	1.03°	**0.99**	2.25	1.01	**0.99**	1.56	1.00	**0.99**
	4Q	1.01	1.01	**0.95**	1.26	1.02	**0.98**	**1.06**	1.05	1.03
	8Q	1.21	0.98	**0.92**	1.37	0.95	**0.92**	**1.15**	1.02	1.05
	12Q	1.22	1.01	**0.92**	1.41	0.95^{**}	**0.93**	1.16	**0.98**	1.02
S.T. Int.	1Q	**0.28**	$5.02^{\circ\circ\circ}$	1.00	0.41	$2.00^{\circ\circ\circ}$	**0.97**	**0.12**	$6.65^{\circ\circ\circ}$	$1.33^{\circ\circ\circ}$
Rate	4Q	1.05	$1.48^{\circ\circ\circ}$	**0.88**	1.36	**1.00**	1.00	**0.31**	$2.74^{\circ\circ\circ}$	$1.18^{\circ\circ}$
	8Q	1.80	$1.16^{\circ\circ\circ}$	**0.82**	2.34	**0.90**	0.95	**0.43**	$2.11^{\circ\circ\circ}$	$1.33^{\circ\circ}$
	12Q	2.17	$1.09^{\circ\circ}$	$\mathbf{0.82}^{*}$	2.81	**0.86**	0.88^{*}	**0.49**	$2.08^{\circ\circ\circ}$	$1.52^{\circ\circ\circ}$
LPS		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.18	-0.04^{***}	$\mathbf{-0.07}^{***}$	2.36	$\mathbf{-0.04}^{***}$	-0.03	2.90	**0.00**	0.02
Growth	4Q	2.90	-0.18^{**}	$\mathbf{-0.34}^{**}$	2.21	$\mathbf{-0.17}^{***}$	-0.11	2.55	$\mathbf{-0.03}$	0.05
	8Q	3.41	-0.31^{**}	$\mathbf{-0.61}^{*}$	2.42	-0.17^{***}	$\mathbf{-0.20}^{**}$	2.58	**0.00**	0.06
	12Q	3.25	-0.23^{***}	$\mathbf{-0.49}^{**}$	2.43	-0.09^{***}	$\mathbf{-0.16}^{**}$	2.52	$\mathbf{-0.01}^{***}$	0.03
Inflation	1Q	1.73	0.00	$\mathbf{-0.03}$	2.35	-0.02	$\mathbf{-0.03}$	1.90	-0.01	$\mathbf{-0.02}$
	4Q	1.45	-0.02	$\mathbf{-0.06}$	1.67	$\mathbf{-0.01}$	0.00	1.55	-0.01	$\mathbf{-0.03}$
	8Q	1.59	-0.03	$\mathbf{-0.03}$	1.82	-0.05	$\mathbf{-0.08}$	1.64	$\mathbf{-0.01}$	0.00
	12Q	1.62	$\mathbf{-0.04}$	-0.02	1.90	-0.06	$\mathbf{-0.10}$	1.60	$\mathbf{-0.03}$	0.02
S.T. Int.	1Q	0.30	$\mathbf{-0.01}^{*}$	$0.50^{\circ\circ\circ}$	0.58	$\mathbf{-0.03}^{***}$	$0.08^{\circ\circ\circ}$	-0.15	**0.00**	$0.29^{\circ\circ\circ}$
Rate	4Q	1.46	$\mathbf{-0.02}^{***}$	0.13	1.79	$\mathbf{-0.14}^{**}$	-0.07	**0.73**	0.01	$0.11^{\circ\circ\circ}$
	8Q	2.04	-0.04^{***}	$\mathbf{-0.11}$	2.58	$\mathbf{-0.31}^{**}$	-0.29	**1.16**	0.03	$0.07^{\circ\circ\circ}$
	12Q	2.25	-0.06^{***}	$\mathbf{-0.17}^{**}$	2.89	-0.44^{**}	$\mathbf{-0.48}^{**}$	**1.38**	0.03	$0.08^{\circ\circ\circ}$

Notes: Forecast evaluation results for the standard VAR with constant coefficients and Minnesota-type priors (CC), the VAR with democratic priors (DP) and the SLM model, for the euro area, the United States and Japan. The top panel displays the RMSE and the bottom panel (minus) the LPS. For the CC model the values of the RMSE and LPS are shown. For the remaining two models the RMSE is divided by the RMSE of the CC model (hence values smaller than one indicate that the CC forecasts are less accurate on average), while for the LPS we subtract the LPS of the CC model (hence the negative values indicate that the CC forecasts are less accurate). The Diebold and Mariano (1995) tests indicate whether the MSE and LPS values of the DP and SLM models are significantly different from the CC model. These take account of the finite sample adjustment of Harvey, Leybourne, and Newbold (1997). For the (few) cases were the estimated variance is negative we use a Bartlett kernel for the variance estimation and then apply the standard Diebold and Mariano (1995) test. We consider one-sided alternative hypothesis. The superscripts *, ** and *** indicate that the DP or SLM model is performing better at significance level 10%, 5% and 1%, respectively. Conversely, the superscript $^{\circ}$ indicates that the CC model is doing significantly better.

Table 1 summarises the results of the forecast evaluations for the three-variable systems for the euro area, the United States and Japan. The top panel reports the accuracy of the point forecasts, while the density forecast evaluations are shown in the bottom panel. For the CC model we display the RMSE and (minus) LPS values[17], for the DP and SLM models these are shown relative to

[17]Therefore lower values indicate better performance. For brevity we will refer to the LPS instead of minus LPS.

Table 2: Forecast evaluation results for 15-variable VARs

RMSE		euro area CC	euro area DP	euro area SLM	United States CC	United States DP	United States SLM	Japan CC	Japan DP	Japan SLM
GDP Growth	1Q	2.15	**0.99***	1.01	2.32	0.98*	**0.97**	4.36	**1.00**	1.00
	4Q	1.96	0.97***	**0.96**	1.99	**0.89****	0.91*	2.78	1.00	**0.95**
	8Q	2.45	0.95**	**0.87***	2.17	**0.90****	0.92	2.70	1.01	**0.97**
	12Q	2.49	0.95**	**0.90****	2.21	**0.95*****	0.95*	2.58	1.02	**0.98**
Inflation	1Q	1.36	0.98	**0.95***	2.20	0.98**	**0.98**	**1.62**	1.04	1.03
	4Q	1.10	0.97	**0.88**	1.28	0.97*	**0.96**	**1.18**	1.01	1.16
	8Q	1.30	0.92*	**0.85**	1.38	0.96**	**0.92**	1.34	**0.94**	1.10
	12Q	1.30	0.92**	**0.86**	1.43	0.95***	**0.93**	1.35	**0.93**	1.06
S.T. Int. Rate	1Q	**0.30**	2.50$^{\circ\circ\circ}$	1.40$^{\circ\circ}$	0.53	1.28$^{\circ\circ}$	**0.90****	0.38	1.79$^{\circ\circ\circ}$	**0.41*****
	4Q	**0.88**	1.31$^{\circ\circ\circ}$	1.23°	1.51	0.95	**0.92**	0.75	1.08	**0.33****
	8Q	**1.47**	1.12$^{\circ\circ\circ}$	1.07	2.37	**0.92***	0.92	0.99	0.94	**0.37*****
	12Q	1.93	1.14$^{\circ\circ\circ}$	**0.93**	2.75	0.90**	**0.90**	1.12	0.93	**0.39*****
LPS		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP Growth	1Q	2.11	**−0.01**	0.00	2.25	**−0.04*****	−0.03	**2.85**	0.01	0.00
	4Q	2.65	−0.05**	**−0.12**	2.22	**−0.20*****	−0.14*	**2.55**	0.02	0.02
	8Q	3.43	−0.20**	**−0.64***	2.41	**−0.20****	−0.18**	2.63	0.02	**−0.01**
	12Q	3.32	−0.16***	**−0.55***	2.43	**−0.11*****	−0.10**	**2.52**	0.01	0.02
Inflation	1Q	1.77	−0.02**	**−0.09***	2.28	−0.01**	**−0.04**	**1.93**	0.01	0.03
	4Q	1.58	−0.04	**−0.18**	1.71	−0.04**	**−0.04**	1.68	**−0.03**	0.10
	8Q	1.70	−0.05	**−0.15**	1.83	−0.06***	**−0.10**	1.80	**−0.12**	0.02
	12Q	1.67	−0.05	**−0.10**	1.88	−0.04	**−0.10**	1.77	**−0.16**	0.02
S.T. Int. Rate	1Q	**0.45**	0.02$^{\circ\circ}$	0.42$^{\circ\circ\circ}$	0.83	**−0.03*****	0.05°	0.53	**−0.07*****	0.34$^{\circ\circ\circ}$
	4Q	**1.33**	0.06°	0.21$^{\circ\circ\circ}$	1.88	−0.09**	**−0.14**	1.15	**−0.12*****	0.07
	8Q	**1.86**	0.09	0.03	2.56	−0.18**	**−0.29**	1.46	**−0.12****	−0.03
	12Q	2.15	0.14°	**−0.11****	2.80	−0.26**	**−0.41***	1.63	**−0.12*****	−0.06**

Notes: See notes to Table 1 for further information.

the CC model. The RMSE values are shown as a ratio compared to the CC model (i.e. values less than one indicate better performance compared to the CC model), while for the LPS values we subtract the LPS of the CC model (therefore, negative values indicate better performance compared to the CC model). Values in bold indicate the best performing model for a given target variable, forecast horizon and evaluation metric. Furthermore, significance of the differences with respect to the CC model is indicated based on one-sided Diebold and Mariano (1995) tests. The asterisk superscripts indicate the significance level in case the DP or SLM model is performing better, while the circle superscripts indicate this when the CC model does better.

For point forecasts of GDP growth and inflation the SLM model is outperforming the CC as well as the DP model for the euro area and the United States. Also for the short-term interest rate the SLM model is performing well in point forecasting for the euro area and the United States. For Japan the story is different, we find that the DP and SLM models are not doing better than the CC model. The main explanation appears to be that for Japan the long-term survey forecasts do

not provide accurate information on long-term developments. The results for the density forecasts are qualitatively similar to those for the point forecasts. For the euro area and the United States the DP and SLM models provide more accurate density forecasts compared to the CC model. The performance of the SLM model relative to the DP model varies depending on the country and variable, with good relative performance for the euro area.

Turning to systems with a larger cross-sectional dimension, Table A2 in Appendix C shows similar patterns for five-variable systems. For the 15-variable VARs Table 2 reports the forecast evaluation results. For the euro area and the United States the SLM model is overall the best performer followed by the DP and the CC model, respectively. For Japan the results are more mixed.

Overall we find that the SLM model is outperforming the CC as well as the DP model in most cases. The better (long-term) forecasting performances for the DP and SLM models over the CC model illustrate the relevance of exploiting the information from the long-term survey forecasts. Moreover, the relevance of allowing for the time variation in the mean is supported by the overall better forecasting performance of the SLM model compared to the DP model.

4.3 Comparison to time-varying parameter models

4.3.1 Benchmark models

Having established the good forecasting performance of the SLM model compared to constant coefficient VARs, the first obvious benchmark is the general time-varying parameter (TVP) VAR, which allows all the coefficients to vary over time:

$$\begin{aligned} y_t &= c_t + \sum_{k=1}^{p} B_{k,t} y_{t-k} + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t), \\ \text{vec}(B_t^c) &= \text{vec}(B_{t-1}^c) + \eta_t, \quad \eta_t \sim N(0, V), \end{aligned} \tag{20}$$

with $B_t^c = [c_t : B_{1,t} : \ldots : B_{p,t}]'$ containing all the VAR coefficients at time t (see e.g. Primiceri, 2005; Cogley and Sargent, 2005). The TVP model is more flexible compared to the SLM model as it allows for time variation both in the unconditional mean and in the impulse response functions.[18] The flexibility comes at the cost of a higher computational burden - the standard TVP-VAR is commonly used for small cross-sectional dimensions - and lack of parsimony. The latter might potentially lead to overfitting and poor forecasting performance. Ultimately, how much flexibility is needed or feasible is an empirical question.

[18] We consider the standard specification of the TVP-VAR, which has a more flexible form of stochastic volatility, since also the matrix A in the decomposition of H_t in Equation (5) is allowed to be time-varying. For more details on the specification of the TVP-VAR model we refer to Primiceri (2005) and D'Agostino, Gambetti, and Giannone (2013).

The second benchmark model we consider is given by a restricted version of the SLM model in which we do not exploit the information from the survey forecasts. Precisely, we drop the second measurement equation given by (4) and obtain the model given by:

$$\begin{aligned} y_t - \psi_t &= \sum_{k=1}^{p} B_k \left(y_{t-k} - \psi_{t-k}\right) + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, H_t\right), \\ \psi_t &= \psi_{t-1} + \eta_t, \quad \eta_t \sim N(0, V_t). \end{aligned} \tag{21}$$

We label this restricted version as the local mean (LM) VAR model. Using the LM model as a benchmark allows to assess the effect of anchoring the local mean by long-term survey forecasts.

Finally, we compare the accuracy of the SLM forecasts of inflation to those from the univariate unobserved components stochastic volatility ($UCSV$) model of Stock and Watson (2007):

$$\begin{aligned} y_t &= \psi_t + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, h_t\right), \\ \psi_t &= \psi_{t-1} + \eta_t, \quad \eta_t \sim N(0, v_t). \end{aligned} \tag{22}$$

The UCSV model is also a restricted version of the SLM (and the LM) model: it is univariate, the VAR coefficients B are set to zero and there is no anchoring of ψ_t to long-term survey forecasts.[19] It has become the standard benchmark for evaluating inflation forecasts (see e.g. Stock and Watson, 2010; Faust and Wright, 2013).

4.3.2 Results

Table 3 reports the relative forecasting performance of the TVP and LM models compared to the SLM model for the three-variable system. Here the SLM model is taken as the base case - in the columns corresponding to the TVP and LM models, values of the RMSE greater than one and positive LPS values indicate better relative performance of the SLM model. The SLM model is outperforming the benchmarks for the euro area and the United States in most cases, both for point and density forecasts. For Japan the SLM model is also performing well with respect to the benchmarks for GDP growth and inflation. For forecasting the Japanese short-term interest rate the TVP and LM models are performing better. Interestingly, the SLM model is often strongly outperforming the more flexible TVP model, which suggests that the time-varying mean might capture the relevant dimension of time variation and/or that the latter model might suffer from over-fitting. The better forecasting performance of the SLM model compared to the LM model, in particular for longer forecast horizons of 8 and 12 quarters, shows the advantage of anchoring the local mean to long-term survey forecasts.

Finally it should be noted that the TVP and LM models are not well able to handle larger cross-sections. Apart from the computational considerations relevant for the TVP model as mentioned

[19] For the UCSV model we use a different prior on the stochastic volatility processes, i.e. we set $\underline{\phi} = 0.2$ in line with Stock and Watson (2007), who instead fix the ϕ parameters at these values.

Table 3: Forecast evaluation results for three-variable VARs

		euro area			United States			Japan		
RMSE		SLM	TVP	LM	SLM	TVP	LM	SLM	TVP	LM
GDP	1Q	2.12	1.05**	**1.00**	**2.57**	1.00	1.03	4.46	**0.97**°	1.06
Growth	4Q	**1.87**	1.19**	1.23	**1.87**	1.10**	1.24*	2.57	**0.99**	1.26
	8Q	**2.14**	1.20***	1.45*	**1.96**	1.13**	1.47**	2.56	**1.00**	1.20
	12Q	**2.27**	1.14***	1.30***	2.07	**1.00**	1.48**	2.49	**1.00**	1.24*
Inflation	1Q	**1.25**	1.11**	1.09**	**2.23**	1.13*	1.05	**1.54**	1.04	1.06*
	4Q	**0.97**	1.19	1.19*	**1.24**	1.48	1.21	**1.09**	1.00	1.13*
	8Q	**1.12**	1.26	1.30*	**1.26**	1.35**	1.16	1.21	**0.99**	1.13
	12Q	**1.12**	1.33**	1.40***	**1.32**	1.42**	1.17***	**1.18**	1.02	1.08
S.T. Int.	1Q	**0.29**	1.01	1.25	0.40	**0.87**°	0.97	0.17	**0.60**°°	0.70°°
Rate	4Q	**0.92**	1.17	1.28	1.36	0.94	**0.94**	0.37	0.70°°°	**0.64**°°°
	8Q	**1.49**	1.34*	1.29*	2.22	1.08	**0.93**	0.58	0.85	**0.64**°°
	12Q	**1.79**	1.42**	1.32**	2.47	1.24*	**0.98**	0.75	0.87	**0.59**°°
LPS		SLM	TVP	LM	SLM	TVP	LM	SLM	TVP	LM
GDP	1Q	2.11	**−0.07**	−0.03	**2.33**	0.02	0.05	**2.92**	0.01	0.01
Growth	4Q	2.56	0.85*	**−0.31**	**2.10**	0.04	0.22	2.60	0.28	**−0.05**
	8Q	2.80	1.96*	**−0.21**	**2.22**	0.10	0.53	2.64	0.47	**−0.04**
	12Q	2.75	1.51*	**−0.15**	**2.28**	0.07	0.68	**2.56**	0.51*	0.03
Inflation	1Q	**1.69**	0.14**	0.07	**2.32**	0.15**	0.09	**1.88**	0.11*	0.05*
	4Q	**1.40**	0.15**	0.21**	**1.68**	0.23*	0.16	**1.52**	0.07	0.11*
	8Q	**1.56**	0.30**	0.45***	**1.74**	0.41*	0.14	**1.64**	0.09	0.11**
	12Q	**1.60**	0.44***	0.62***	**1.80**	0.51**	0.20	**1.62**	0.12**	0.16***
S.T. Int.	1Q	0.80	**−0.83**°°°	−0.15°°°	0.66	**−0.50**°°°	0.26***	0.14	**−0.98**°°°	−0.05°°°
Rate	4Q	1.60	**−0.16**	0.03	1.71	**−0.07**	0.07	0.83	**−0.48**°°°	−0.13°°°
	8Q	**1.94**	0.38	0.16	2.30	0.11	**−0.11**	1.23	**−0.22**°°	−0.17°°°
	12Q	**2.08**	0.48**	0.21**	2.41	0.31**	**−0.04**	1.46	−0.11	**−0.20**°°°

Notes: Forecast evaluation results for the SLM, TVP and LM models for three-variable systems. The RMSE and LPS are reported in absolute values for the SLM model and relative to the SLM model for the remaining two models. The superscripts * denote significantly better performance of the SLM model, while the superscripts ° denote that the TVP or the LM model is doing significantly better. See notes Table 1 for further information.

earlier, also shrinkage of the coefficients is important to control for overfitting in higher dimensional VARs. For the LM model the anchoring of the local mean by off-model information, as in the SLM model, becomes even more crucial for higher dimensional VARs.

Table 4 shows the point and density forecasting accuracy of the SLM and UCSV models for inflation. For the United States and Japan the SLM model provides better point and density forecasts. For the euro area instead there is not much difference in point forecast accuracy, except for three years ahead, when the SLM model is doing better. The improvement of the SLM model over the UCSV model for the euro area lies in the long-term density forecasts. To shed more light on the sources of those differences, Figure A1 in Appendix C reports the 12-quarter-ahead density forecasts for the SLM model (left panel) and the UCSV model (right panel) together with the data realisations. It turns out that forecasts from the UCSV model are more volatile between

Table 4: Forecast evaluation results for the SLM and the UCSV models

RMSE		euro area SLM	euro area UCSV	United States SLM	United States UCSV	Japan SLM	Japan UCSV
Inflation	1Q	**1.25**	1.03	**2.23**	1.08	**1.54**	1.05
	4Q	0.97	**0.98**	**1.24**	1.26	**1.09**	1.12
	8Q	1.12	**0.99**	**1.26**	1.17	**1.21**	1.12
	12Q	**1.12**	1.06	**1.32**	1.18^{***}	**1.18**	1.09^{*}
LPS		SLM	UCSV	SLM	UCSV	SLM	UCSV
Inflation	1Q	**1.69**	0.02	**2.32**	0.01	**1.88**	0.03
	4Q	1.40	**−0.01**	**1.68**	0.27^{**}	**1.52**	0.11
	8Q	**1.56**	0.13	**1.74**	0.31	**1.64**	0.17^{*}
	12Q	**1.60**	0.33^{**}	**1.80**	0.57^{**}	**1.62**	0.33^{***}

Notes: Forecast evaluations for the three-variable SLM model and the univariate UCSV model for inflation. The SLM model is taken as the base case here. The superscripts * denote significantly better performance of the SLM model, while the superscripts ° indicate that the UCSV model is doing significantly better. See notes Table 1 for further information.

the updates and have become much more dispersed following the financial crisis.

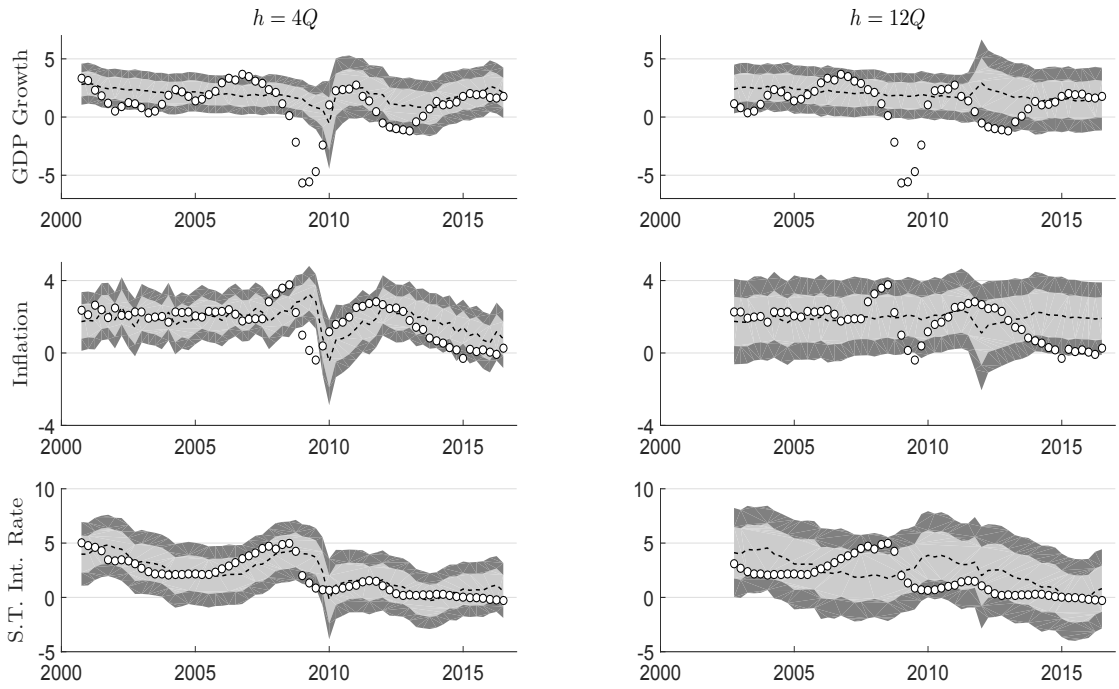
5 More detailed analysis for the euro area

To highlight the relevant features of the SLM model (also relative to the other approaches) this section provides more detailed illustrations focusing on the euro area. Precisely, we look at the properties of the density forecasts of the SLM model, we investigate more in detail the performance of the model relative to the CC model, we analyse the posterior distributions of the end-of-sample unconditional and local means and of the impulse response functions from different models and finally we construct conditional forecasts from the SLM model based on different scenarios for the long-term inflation expectations.

5.1 Calibration of density forecasts

Figure 4 shows the density forecasts of the SLM model together with the data realisations. For brevity we only focus on forecast horizons of 4 and 12 quarters. For GDP growth the density forecasts typically include the realisation with the exception of the financial crisis period when the outcomes were much lower than the forecasts. After the crisis the coverage bands widened considerably, in particular for the longer forecast horizon. For inflation the realisations are mostly within the coverage bands and the effect of the crisis on inflation forecasts is less pronounced than for GDP growth. What is more striking is the systematic over-prediction of inflation in the recent years. This is due to the persistently low inflation rates combined with the relatively high survey

Figure 4: Density forecasts for the SLM model



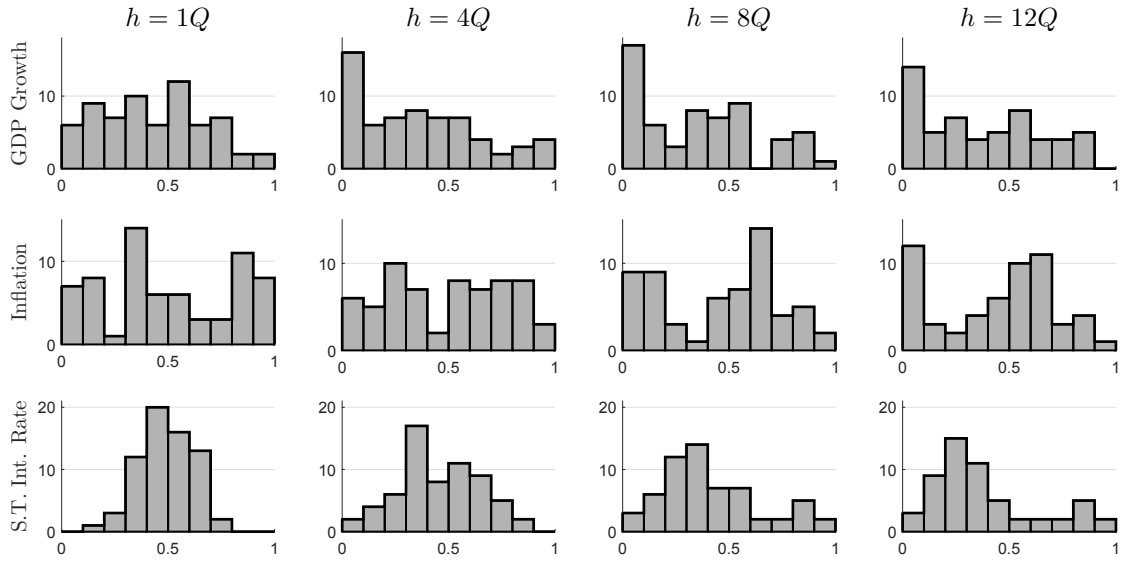
Notes: Predictive distributions for the SLM model with three variables for the euro area for forecast horizons of 4 (left panel) and 12 quarters (right panel). The forecasts and realisations correspond to average rates over the first (for 4-quarter ahead forecasts) and third (for 12-quarter ahead forecasts) year. The x-axes indicate the time of the realisations. The density forecasts are summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. Realisations are indicated by white circles.

forecasts for inflation in this period (which push the SLM model forecasts up). Lastly, for the interest rate the predictive coverage bands include the observations but are relatively wide.

To investigate the calibration of the predictive densities more formally we look at the probability integral transforms (PITs). The PITs are the values of the predictive cumulative distribution function (CDF) evaluated at the out-of-sample realisations and for well calibrated density forecasts they should follow a uniform distribution on the interval [0 1].[20] Given that our out-of-sample period is relatively short we look at the histograms of the PITs instead of doing formal tests. Figure 5 shows the histograms for the three variables of interest and the four forecast horizons. For GDP growth and inflation the density forecasts seem to reflect the distribution of the outcomes reasonably well. For longer forecast horizons perhaps too many realisations fall in the left tail of the predictive distribution, suggesting some over-prediction. For the short-term interest rate the predictive distributions seem too wide, in particular for the shorter forecast horizons. For the longer horizons there is again some evidence of over-prediction.

[20]From the set of simulated forecasts we obtain the PIT value as the proportion of simulated forecasts falling below the realisation.

Figure 5: Probability integral transforms for the SLM model



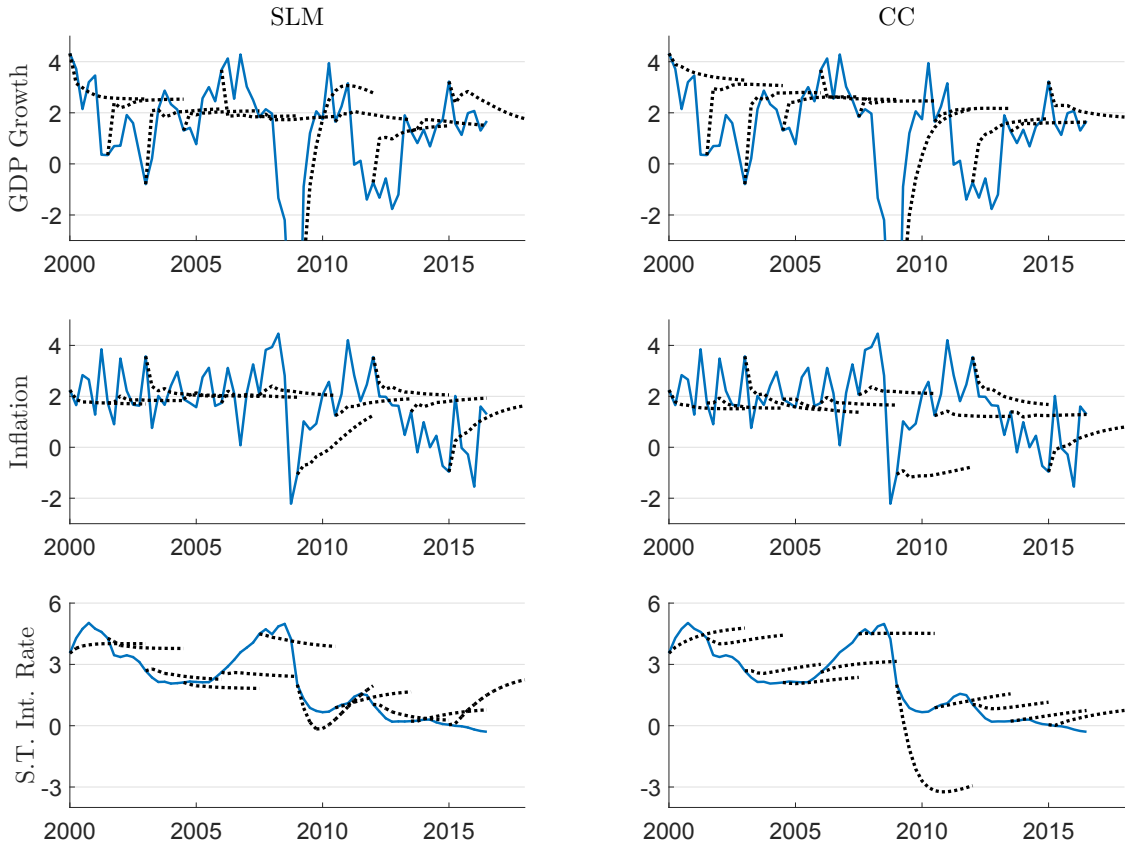
Notes: The histograms show the PITs for the SLM model with three variables for the euro area. The columns correspond to the forecast horizons and the rows to the variables.

5.2 Comparison to the CC model

We first look at the forecast *paths*, commonly referred to as "tentacle" plots, for the target variables at different points in the evaluation sample, see Figure 6. For GDP growth the forecasts from the CC model initially converge to somewhat too high values, as the unconditional mean is estimated on the sample dominated by the early period of higher growth rates. For the SLM model this is not the case, illustrating the advantage of including off-model information in the presence of structural changes. For inflation, we can observe some instability in the unconditional mean in the CC model, as the crisis observations influence the long-term forecasts to an implausible extent (we return to this issue in the next section). This shows the advantage of using survey forecasts to better anchor the unconditional or local mean and consequently the long-term forecasts. On the other hand, using the information from long-term survey forecasts appears to exacerbate the inflation over-prediction problem in the recent years (the CC forecasts are lower than the SLM forecasts). For the short-term interest rate forecasts we can again observe the instability of the unconditional mean estimates in the CC model as a few additional observations can result in a fundamentally different long-term forecast. These patterns can also be observed in the density forecasts, see Figure A2. In particular, compared to the SLM model, the PIT histograms for the CC model (see Figure A3) exhibit more outcomes in the left tail of the predictive distribution for GDP growth and less for inflation (consistent with the stronger over-prediction of GDP by the CC model and of inflation by the SLM model).

To investigate how the relative performance of the SLM and CC models evolves over time we look

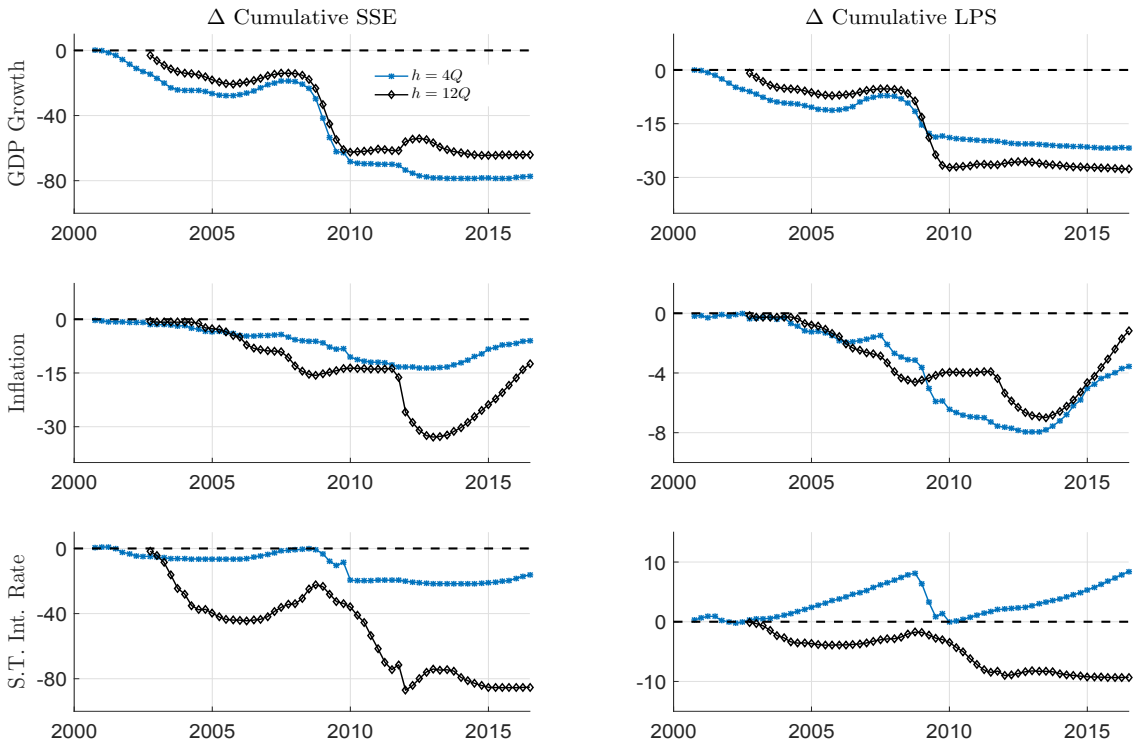
Figure 6: Forecast paths for the SLM and CC models



Notes: Forecast paths (up to 12 quarters ahead) for the SLM (left panels) and CC (right panels) models with three variables for the euro area. The first forecast path originates in 2000Q1; subsequently, for better readability, we show forecast paths with origins six quarters apart.

at the cumulative sums of the squared forecast errors for point forecasts and the log predictive score values for the density forecasts, see Figure 7. We subtract the cumulative sum for the CC model from that of the SLM model. Hence, declining cumulative sums indicate better performance of the SLM model in that period, and vice versa for upward sloping lines. For brevity we only consider forecast horizons of 4 and 12 quarters. In line with the results shown in Figure 6, we see that for GDP growth the SLM model performs better than the CC model for the period until and including the financial crisis. Thereafter, the accuracy of both models is comparable. For inflation, the better performance of the SLM model persists till around 2013. Thereafter, the relative accuracy of the SLM model deteriorates, reflecting the over-prediction issues mentioned above. For the short-term interest rate the picture is a bit mixed across horizons and accuracy measures. The SLM model mostly outperforms the CC model till the recent period where both models appear to have comparable accuracy. The exception is the density forecasts 4 quarters ahead where the CC model does better.

Figure 7: Relative forecast performance of the SLM model compared to the CC model



Notes: Differences of the cumulative sums over time of squared forecast errors (left panels) and of LPS values (right panels) between the SLM and the CC models with three variables for the euro area. The cumulative sum for the CC model is subtracted from that of the SLM model; therefore, declining values indicate a better performance of the SLM model. Blue and black lines correspond to the forecast horizons of 4 and 12 quarters, respectively. The x-axes indicate the time of the realisations.

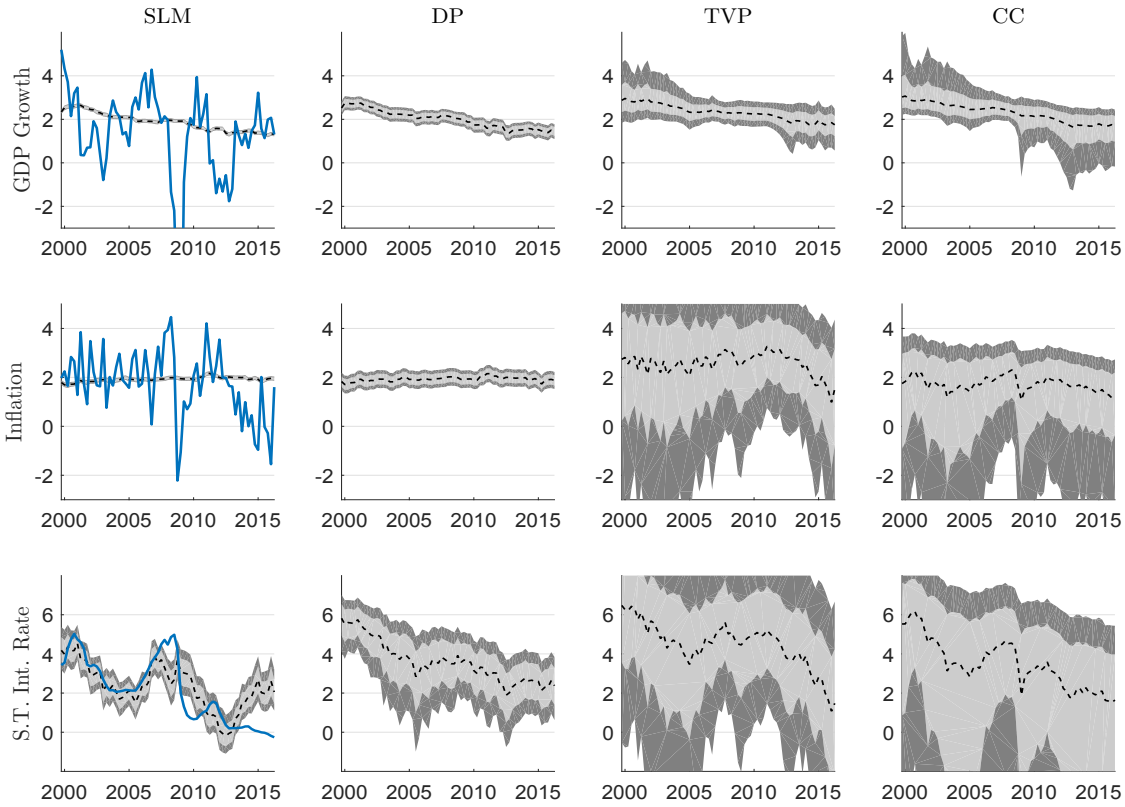
Overall, the long-term survey forecasts appear to carry useful (forward-looking) off-model information, with the exception of inflation in the most recent period, and help to make the forecast revisions less volatile.

5.3 Posterior distribution of the local mean

The posterior distribution of the end-of-sample local or unconditional mean is highly relevant for forecasting as it determines the end point of the forecast path.[21] In this section we analyse the evolution of the posterior distribution of the end-of-sample local or unconditional mean for the SLM, DP, TVP and CC models. Precisely, for each point s in the forecast evaluation sample we report the posterior distribution of the local or unconditional mean at time $s-1$, with the

[21] For time invariant $B_1, \ldots, B_p$ satisfying the stability condition, the point forecast will converge towards the local or unconditional mean, see also Equation (3).

Figure 8: Posterior distributions for the end-of-sample local and unconditional means



Notes: Posterior distributions of the end-of-sample local and unconditional means for the SLM, DP, TVP and CC models with three variables for the euro area. These posterior distributions are obtained recursively in the out-of-sample forecast evaluations. The x-axes indicate the end of the estimation sample (the forecast origin). The posterior distributions are summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. The blue lines in the first column correspond to the data.

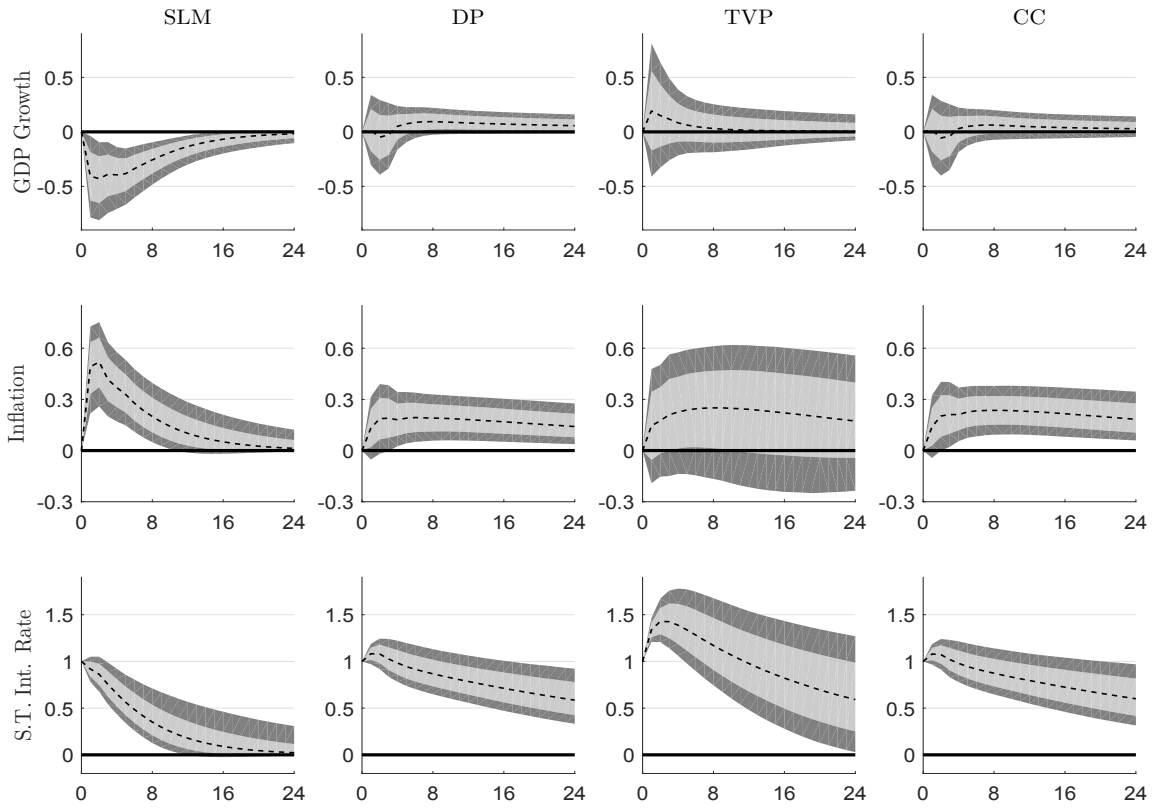
parameters estimated over $t = 1, \ldots, s-1$. For the SLM and DP models the local and unconditional means, respectively, are sampled directly. For the CC model we obtain the draws of the unconditional mean $\mu^{(i)}$ from

$$\mu^{(i)} = (I - B_1^{(i)} - \ldots - B_p^{(i)})^{-1} c^{(i)}, \tag{23}$$

for each of the Gibbs draws, i, of B and c. We proceed analogously for the TVP model, where we take the end-of-sample values B_{s-1}^c.

Figure 8 shows the end-of-sample posterior distributions over 1999Q4-2016Q2. For the SLM model the posterior distribution is tightly concentrated around the median for GDP growth and inflation and changes little over time. For the short-term interest rate the posterior distribution is somewhat more disperse, as there is no anchoring through the survey forecasts. These findings are in line with the results on the in-sample posterior distribution of the local mean shown in

Figure 9: Impulse response functions



Notes: Impulse response functions to a 100 basis points increase in the euro area short-term interest rate for VARs with three variables estimated over the full sample. For the TVP model the IRFs are based on the end-of-sample parameters. The posterior distribution is summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands.

Figure 2. For the DP model the posterior distributions are somewhat wider but we still see a strong anchoring for GDP growth and inflation.[22] By contrast, the distributions for the TVP and CC models show much more dispersion and variation over time, especially for inflation and the short-term interest rate. This explains the occasional occurrence of considerable changes between consecutive forecast updates for the CC model reported in Figure 6.

5.4 Impulse response functions

A common comparison between different VAR models is to look at the impulse response functions (IRFs) to an identified monetary policy shock. In this section we consider IRFs to a 100 basis

[22]We have also experimented with a prior variance of 0.01 on the unconditional mean. This yields a stronger anchoring but in terms of forecast accuracy there was no improvement.

points shock to the short-term interest rate, identified via a Cholesky scheme with the interest rate ordered as last in the VAR.

Figure 9 displays the IRFs for the SLM, DP, TVP and CC models based on the parameters estimated over the full sample for the three-variable VAR for the euro area.[23] For the TVP model the IRFs are based on the end-of-sample parameters, i.e. they are representative for 2016Q3. For the SLM model the IRFs go to 0 after around 3 to 4 years. By contrast, in case of the CC and DP models, the IRFs for inflation and interest rate are significantly different from 0 even after 6 years. This is contrary to what would be expected for a temporary shock and highlights the problematic assumption of a constant unconditional mean in those models, which translates into high persistence in the coefficients B. For the TVP model, the median responses are also relatively persistent, but the coverage bands are much wider and as a result the responses are not significantly different from 0 for inflation. As for GDP growth, whereas the IRFs from the SLM model are negative and significant, those from the DP, TVP and CC models are essentially equal to 0. Finally, all the models suffer from the so called price puzzle, i.e. inflation rises in response to monetary policy tightening. This is a common problem in small VARs and can be solved by enlarging the dataset, see e.g. Bańbura, Giannone, and Reichlin (2010).

5.5 Scenarios for long-term expectations

The forecasting exercises considered so far did not incorporate any information on the possible future evolution of survey forecasts. In this subsection we extend this setting by constructing forecasts *conditional* on scenarios for the long-term expectations of inflation.

We obtain such conditional forecasts following the ideas of data augmentation, i.e. by treating the out-of-sample forecasts as latent variables and sampling them alongside the model parameters in an (additional) Gibbs sampling step. Let $z^*_{T+1}, \dots, z^*_{T+h}$ denote an h step-ahead scenario and let Z^* contain the in-sample survey forecasts as well as the scenario. Similarly, let Y^* contain the in-sample observations (i.e. $y_1, \dots, y_T$) as well as the part we want to forecast (i.e. $y^*_{T+1}, \dots, y^*_{T+h}$). Given (some initial) values of the parameters we can sample $y^*_{T+1}, \dots, y^*_{T+h}$ from

$$p\left(y^*_{T+1}, \dots, y^*_{T+h} \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T, \{Y, Z^*\}\right), \tag{24}$$

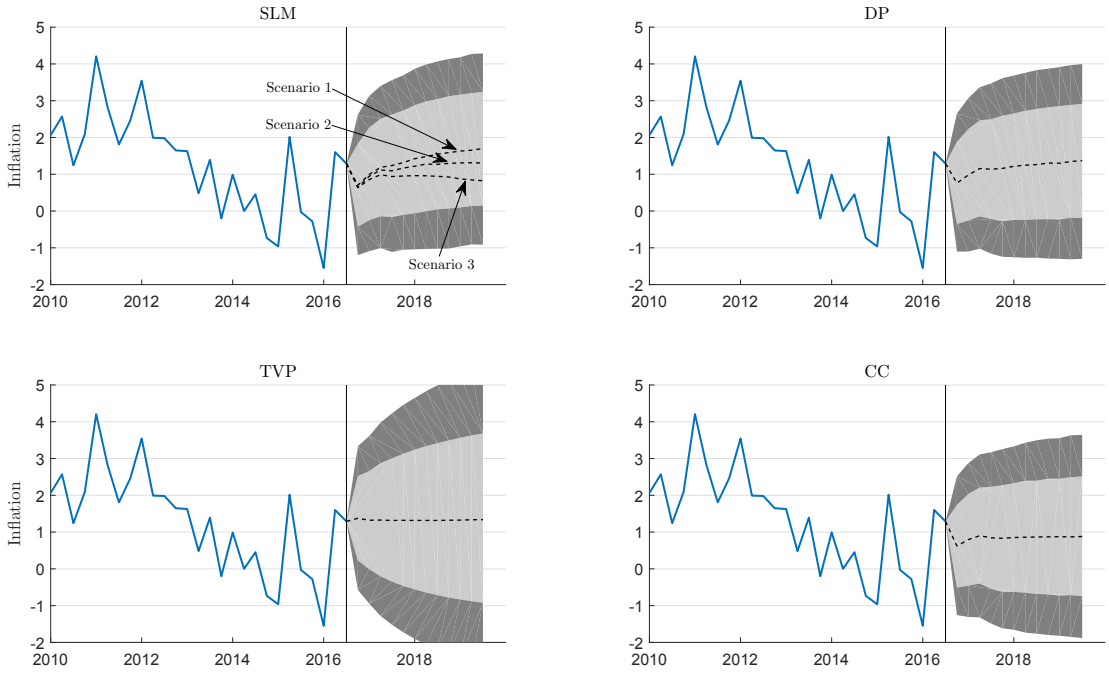
by iteratively sampling h times from the VAR equation in (1).[24] Then conditional on values for $\{Y^*, Z^*\}$ we can use the usual Gibbs sampling steps to sample the parameters.[25]

[23] We also considered the same setting for the five variable dataset and found no significant differences.

[24] These values are sampled at step 6 of the algorithm explained in Appendix A (the previous step 6 becomes step 7).

[25] With the difference that the time series is now of length $T + h$ instead of T. For reasonable scenarios the approach does not lead to a distortion of the posterior distribution of the parameters. For example, Figure A4 in the Appendix shows that forecasts conditional on the long-term survey forecasts remaining constant are very similar to the pure out-of-sample forecasts and we also find only small differences in the posterior distribution of the parameters.

Figure 10: (Conditional) forecasts for inflation



Notes: 12-quarter-ahead forecasts for euro area inflation for the SLM, DP, TVP and CC models with three variables. The in-sample period ends at 2016Q3 and the forecast period is 2016Q4 to 2019Q3. The forecasts for the SLM model are conditional on different scenarios for future developments in the long-term survey forecasts: 1 - kept at the value of 2016Q3; 2 - linearly declining to 1.5% in 2019Q3; 3 - linearly declining to 1% in 2019Q3. For the DP, TVP and CC models the forecasts are unconditional. The forecast distribution is summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. For the SLM model, the distribution is shown for scenario 1, for the remaining scenarios only the median is reported.

As an illustration we look at inflation forecasts conditional on three scenarios for the long-term inflation expectations for the three-variable VAR for the euro area. Precisely, we take the full data sample (till 2016Q3) and we forecast 12 quarters ahead conditional on the following assumptions for the long-term survey forecasts of inflation (over 2016Q4-2019Q3):

Scenario 1 The forecasts are held constant at the value of 2016Q3;

Scenario 2 The forecasts decline linearly over 12 quarters, reaching 1.5% in 2019Q3;

Scenario 3 The forecasts decline linearly over 12 quarters, reaching 1.0% in 2019Q3.

For GDP growth, the long-term survey forecasts are held constant at the value of 2016Q3 in all the scenarios.

The resulting forecasts for inflation are shown in the upper left panel of Figure 10. For scenario 1 we display the median together with the 68% and 90% coverage bands, whereas only the median is reported for scenarios 2 and 3 for better readability. The forecasts for scenarios with declining

expectations are clearly lower, showing the relevance of the survey forecasts for the SLM model forecasts. For comparison, the figure also shows the *unconditional* forecasts from the DP, TVP and CC models. The forecasts from the TVP and CC models are relatively low, close to those from the SLM model with scenario 3, as it appears that the local and unconditional means in these models are affected by the low inflation outcomes in recent years.[26]

6 Sensitivity analysis

This section reports a range of robustness checks for the forecast evaluations, whereby we vary a number of implementation choices. We take as benchmarks the constant coefficient three-variable VARs considered in Section 4.2. The tables we refer to below can be found in Appendix C.

Rolling window estimation

In the main forecasting exercises an expanding estimation window was used. However, in the presence of time variation in the data generating process, rolling estimation windows might result in better forecasting performance, especially for constant parameter models, because the early part of the estimation sample might distort the parameter estimates. Table A3 reports the results of forecast evaluations for the CC, DP and SLM models for a rolling estimation window of 80 quarters. We find that also with this implementation the SLM model comes out best.[27]

Constant parameters out-of-sample

To generate forecasts in the main setup we simulated the out-of-sample paths of the time-varying parameters. For the CC and DP models this amounts to simulating the stochastic volatility processes for the VAR innovations for $t = T+1, \ldots, T+h$. For the SLM model also the stochastic volatility process in the random walk for the local mean and the local mean process are simulated. Table A4 compares the forecasting accuracy of the three models for the case when the values of the time-varying parameters at $t = T+1, \ldots, T+h$ are kept at the sampled values at $t = T$. This setup does not change the conclusion that the SLM model provides the most accurate forecasts.

Constant variance

Table A5 compares the forecasting accuracy under the assumption of no time variation in the

[26] As the local mean (trend) in the UCSV model tends to follow actual inflation quite closely, the unconditional forecasts from this model starting in 2016Q3 are the lowest among the models considered. They are also characterised by the widest coverage bands, see Figure A5.

[27] Qualitatively similar results are obtained for rolling estimation windows of 40 and 60 quarters. Interestingly, the SLM model performs best for the expanding estimation window. It appears that once we account for time variation in the mean, the early part of the sample provides useful information. The results are available upon request.

variances of the innovations.[28] Also in this case the SLM model provides the best forecasts.[29]

Robust evaluation criteria

The forecast evaluation criteria we considered, the RMSE and the LPS based on a fitted Gaussian distribution, might be disproportionately affected by a few large forecast errors. Table A6 reports the more robust mean absolute errors (MAE) for the point forecasts, but also LPS values based on a fitted standardised t-distribution and the continuous ranked probability score (CRPS) (Gneiting and Raftery, 2007), for the density forecasts. These robust evaluation criteria lead to the same conclusion that the SLM model is overall the best performer.

Comparison to pure survey forecasts

Previous studies such as Faust and Wright (2009), Faust and Wright (2013) and Aiolfi, Capistrán, and Timmermann (2011) show that survey forecasts by themselves provide excellent point forecasts compared to model-based forecasts. Table A7 confirms these findings for our evaluation sample. Compared to the long-term Consensus forecasts, the SLM model forecasts result in some modest improvements for short horizons for the euro area and the United States and more sizable improvements for Japan (where as mentioned above, survey forecasts do not track so well the long-term developments in the variables of interest). For longer horizons survey forecasts are overall comparable in terms of accuracy to those from the SLM model for the euro area and the United States. This is not so surprising, given that survey forecasts provide an "anchor" for the model-based forecasts. However, even when model-based forecasts do not (significantly) improve upon survey forecasts, they offer many features not offered by surveys, such as predictive distributions for a variety of forecast horizons, scenarios and also flexibility in terms of target variables.

7 Conclusion

The focus of this paper is how to model time variation in VAR models in a forecasting context. Precisely, we propose a local mean VAR model which allows for time variation in the unconditional mean, but keeps the VAR coefficients constant over time. Furthermore, we anchor the local mean to long-term Consensus forecasts (off-model information). From a computational point of view, the model can be used for relatively large cross-sections.

Through out-of-sample forecasting exercises we show that the local mean VAR model has better

[28] We implement this by putting a very tight prior on $\phi_{J,i}$, $\forall \{J, i\}$ in (14)-(16), this results in constant $\sigma^2_{J,i,t}$ over time.

[29] In fact the differences in forecasting accuracy between the model versions with stochastic volatility and with constant variance are very small, suggesting that this element might not be crucial for the evaluation sample considered in this paper.

forecasting performance compared to constant coefficient VARs and general time-varying parameter VARs. The off-model information can provide a useful signal on the "endpoints" and, in particular, reduces the uncertainty around them. The proposed framework not only results in better forecasting performance but also in more reasonable impulse response functions. It can be also used for scenario analysis, considering e.g. different future paths for long-term expectations.

The forecasting performance of the model studied here is strongly linked to how well the off-model information can capture the low frequency movements in the target variables. Long-term survey forecasts provided by Consensus Economics seem to offer a good "anchor" for the main macroeconomic variables for the euro area and the United States. For Japan and more recently for inflation in the euro area they appear to provide a less useful signal. Other types of off-model information, e.g. potential output growth estimates or short-term inflation expectations, could also be considered. A systematic evaluation of such anchors is left for future research.

References

Aastveit, K. A., A. Carriero, T. E. Clark, and M. Marcellino (2017): "Have Standard VARS Remained Stable Since the Crisis?," *Journal of Applied Econometrics*, 32(5), 931–951.

Aiolfi, M., C. Capistrán, and A. G. Timmermann (2011): "Forecast combinations," in *Oxford Handbook on Economic Forecasting*, ed. by M. P. Clements, and D. F. Hendry, pp. 355–388. Oxford University Press.

Andrle, M., and J. Bruha (2017): "Forecasting and policy analysis with trend-cycle Bayesian VARs," Discussion paper, unpublished.

Ang, A., G. Bekaert, and M. Wei (2007): "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 54(4), 1163–1212.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017): "Tracking the Slowdown in Long-Run GDP Growth," *The Review of Economics and Statistics*, 99(2), 343–356.

Bańbura, M., D. Giannone, and L. Reichlin (2010): "Large Bayesian vector auto regressions," *Journal of Applied Econometrics*, 25(1), 71–92.

Barnett, A., H. Mumtaz, and K. Theodoridis (2014): "Forecasting UK GDP growth and inflation under structural change. A comparison of models with time-varying parameters," *International Journal of Forecasting*, 30(1), 129 – 143.

Belmonte, M. A., G. Koop, and D. Korobilis (2014): "Hierarchical Shrinkage in Time-Varying Parameter Models," *Journal of Forecasting*, 33(1), 80–94.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005): "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach," *The Quarterly journal of economics*, 120(1), 387–422.

Carriero, A., T. E. Clark, and M. Marcellino (2016a): "Common drifting volatility in large Bayesian VARs," *Journal of Business & Economic Statistics*, 34(3), 375–390.

——— (2016b): "Large vector autoregressions with stochastic volatility and flexible priors," Working Paper 16-17, Federal Reserve Bank of Cleveland.

Carter, C. K., and R. Kohn (1994): "On Gibbs sampling for state space models," *Biometrika*, 81(3), 541–553.

Chan, J. C., T. E. Clark, and G. Koop (2017): "A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations," *Journal of Money, Credit and Banking*, 50(1), 5–53.

Clark, T. E. (2011): "Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility," *Journal of Business & Economic Statistics*, 29(3), 327–341.

Clark, T. E., and T. Doh (2014): "Evaluating alternative models of trend inflation," *International Journal of Forecasting*, 30(3), 426–448.

Clark, T. E., and M. W. McCracken (2006): "Forecasting of small macroeconomic VARs in the presence of instabilities," Research Working Paper RWP 06-09, Federal Reserve Bank of Kansas City.

Clark, T. E., and F. Ravazzolo (2015): "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics*, 30(4), 551–575.

Cogley, T., and T. J. Sargent (2001): "Evolving Post-World War II U.S. Inflation Dynamics," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and K. Rogoff, vol. 16, p. 331373.

Cogley, T., and T. J. Sargent (2005): "Drifts and volatilities: monetary policies and outcomes in the post WWII US," *Review of economic dynamics*, 8(2), 262–302.

Crump, R. K., S. Eusepi, and E. Moench (2016): "The term structure of expectations and bond yields," Staff Reports 775, Federal Reserve Bank of New York.

D'Agostino, A., L. Gambetti, and D. Giannone (2013): "Macroeconomic forecasting and structural change," *Journal of Applied Econometrics*, 28(1), 82–101.

de Wind, J., and L. Gambetti (2014): "Reduced-rank time-varying vector autoregressions," CPB Discussion Paper 270, CPB Netherlands Bureau for Economic Policy Analysis.

Del Negro, M., and G. E. Primiceri (2015): "Time varying structural vector autoregressions and monetary policy: a corrigendum," *The review of economic studies*, 82(4), 1342–1345.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Dijk, D., S. J. Koopman, M. Wel, and J. H. Wright (2014): "Forecasting interest rates with shifting endpoints," *Journal of Applied Econometrics*, 29(5), 693–712.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and conditional projection using realistic prior distributions," *Econometric reviews*, 3(1), 1–100.

Durbin, J., and S. J. Koopman (2012): *Time series analysis by state space methods*, vol. 38. OUP Oxford.

Eisenstat, E., J. C. C. Chan, and R. W. Strachan (2016): "Stochastic Model Specification Search for Time-Varying Parameter VARs," *Econometric Reviews*, 35(8-10), 1638–1665.

Fagan, G., J. Henry, and R. Mestre (2001): "An area-wide model (AWM) for the euro area," Working Paper Series 0042, European Central Bank.

Faust, J., and J. H. Wright (2009): "Comparing Greenbook and reduced form forecasts using a large realtime dataset," *Journal of Business & Economic Statistics*, 27(4), 468–479.

——— (2013): "Forecasting inflation," in *Handbook of economic forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2A, pp. 3–56. Elsevier.

Frühwirth-Schnatter, S., and H. Wagner (2010): "Stochastic model specification search for Gaussian and partial non-Gaussian state space models," *Journal of Econometrics*, 154(1), 85–100.

Garnier, C., E. Mertens, and E. Nelson (2015): "Trend Inflation in Advanced Economies," *International Journal of Central Banking*, 11(S1), 65–136.

Gneiting, T., and A. E. Raftery (2007): "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, 102(477), 359–378.

Götz, T., and K. Hauzenberger (2017): "Large Mixed-Frequency VARs with a Parsimonious Time-Varying Parameter Structure," unpublished.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the equality of prediction mean squared errors," *International Journal of forecasting*, 13(2), 281–291.

Kadiyala, K. R., and S. Karlsson (1997): "Numerical methods for estimation and inference in Bayesian VAR-models," *Journal of Applied Econometrics*, pp. 99–132.

Karlsson, S. (2013): "Forecasting with Bayesian vector autoregressions," in *Handbook of Economic Forecasting, Volume 2B*, ed. by G. Elliott, and T. Timmermann. North Holland, Elsevier.

Kim, C.-J., and C. R. Nelson (1999): *State-Space Models with Regime Switching*. MIT Press.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic volatility: likelihood inference and comparison with ARCH models," *The review of economic studies*, 65(3), 361–393.

Koop, G., and D. Korobilis (2013): "Large time-varying parameter VARs," *Journal of Econometrics*, 177(2), 185–198.

Kozicki, S., and P. A. Tinsley (2012): "Effective Use of Survey Information in Estimating the Evolution of Expected Inflation," *Journal of Money, Credit and Banking*, 44(1), 145–169.

Lenza, M., and M. Jarociński (2016): "An inflation-predicting measure of the output gap in the euro area," Working Paper Series 1966, European Central Bank.

Mertens, E. (2016): "Measuring the Level and Uncertainty of Trend Inflation," *The Review of Economics and Statistics*, 98(5), 950–967.

Primiceri, G. E. (2005): "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies*, 72(3), 821–852.

Sims, C. A. (2001): "Comment on Sargent and Cogleys Evolving post World War II US inflation dynamics," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and K. Rogoff, vol. 16, p. 373379.

Stock, J. H. (2001): "Discussion of Cogley and Sargent Evolving post World War II US inflation dynamics," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and K. Rogoff, vol. 16, p. 379387.

Stock, J. H., and M. W. Watson (2007): "Why has US inflation become harder to forecast?," *Journal of Money, Credit and banking*, 39(s1), 3–33.

Stock, J. H., and M. W. Watson (2010): "Modeling inflation after the crisis," *Proceedings - Economic Policy Symposium - Jackson Hole*, pp. 173–220.

Stock, J. H., and M. W. Watson (2012): "Disentangling the Channels of the 2007-2009 Recession," *Brookings Papers on Economic Activity*, pp. 81–135.

Villani, M. (2009): "Steady-state priors for vector autoregressions," *Journal of Applied Econometrics*, 24(4), 630–650.

Wright, J. H. (2013): "Evaluating Real-Time Var Forecasts With An Informative Democratic Prior," *Journal of Applied Econometrics*, 28(5), 762–776.

Appendices

A Details on the Gibbs sampler

We obtain draws from the posterior distribution of $\{\{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T\}$ via a Gibbs sampling scheme.

Let $\{Y, Z\}$ denote the full set of observations on the variables (Y) and the survey forecasts (Z). Further, let $S^T = \{s_H^T, s_V^T, s_G^T\}$ denote the mixture states in the algorithm of Kim, Shephard, and Chib (1998) (from now on referred to as the KSC algorithm). We incorporate the corrigendum of Del Negro and Primiceri (2015) and sample the mixture states S^T in the last step of the Gibbs sampling scheme.

Given a set of starting values for the parameters, the Gibbs sampler proceeds by drawing recursively from the following conditional distributions:

1. $p\left(\{\Lambda^T, V^T, G^T\} \mid \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T, S^T, \{Y, Z\}\right)$
2. $p\left(\{\phi_H, \phi_V, \phi_G\} \mid \{\Lambda^T, V^T, G^T\}, A, B, \psi^T, S^T, \{Y, Z\}\right)$
3. $p\left(A \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, B, \psi^T, S^T, \{Y, Z\}\right)$
4. $p\left(B \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, \psi^T, S^T, \{Y, Z\}\right)$
5. $p\left(\psi^T \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, S^T, \{Y, Z\}\right)$
6. $p\left(S^T \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T, \{Y, Z\}\right)$

We next explain these steps in detail.

1. Sampling from $p\left(\{\Lambda^T, V^T, G^T\} \mid \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T, S^T, \{Y, Z\}\right)$

Part a: Λ^T

Consider the innovations ε_t from the VAR equation at time t

$$\varepsilon_t = (y_t - \psi_t) - \sum_{i=1}^{p} B_i \left(y_{t-i} - \psi_{t-i}\right), \quad t = p+1, \ldots, T. \tag{25}$$

Conditional on B and ψ^T the series of ε_t is observable. Recall that $\varepsilon_t \sim N\left(0, A^{-1}\Lambda_t \left(A^{-1}\right)'\right)$ where $\Lambda_t = \text{diag}\left(\sigma^2_{H,1,t}, \ldots, \sigma^2_{H,M,t}\right)$. Hence, the standardised innovations $\tilde{\varepsilon}_t = A\varepsilon_t$ have a theoretical variance of Λ_t. We next show how we use the KSC algorithm to sample from the

posterior distributions of the diagonal elements of Λ_t for $t = p+1, \ldots, T$. For $i = 1, \ldots, M$ we square the series of $\tilde{\varepsilon}_{i,p+1}, \ldots, \tilde{\varepsilon}_{i,T}$, take logarithms and obtain the following linear state space model:

$$\ln\left(\tilde{\varepsilon}_{i,t}^2 + 0.001\right) = \ln\left(\sigma_{H,i,t}^2\right) + \ln\left(\gamma_{H,i,t}^2\right), \tag{26}$$

$$\ln \sigma_{H,i,t+1}^2 = \ln \sigma_{H,i,t}^2 + u_{H,i,t}, \tag{27}$$

where $\ln\left(\gamma_{H,i,t}^2\right)$ follows a log $\chi^2(1)$ distribution. As in the KSC algorithm we obtain an approximating linear Gaussian state space model by replacing $\ln\left(\gamma_{H,i,t}^2\right)$ with a seven state Gaussian mixture approximation, with the state-dependent means and variances given by $(m_j - 1.2704)$ and v_j^2 respectively (for further details we refer to Kim, Shephard, and Chib, 1998). Let $s_{H,i,t} \in \{1, \ldots, 7\}$ denote the state, then conditional on all the states s_H^T we have a linear Gaussian state space model. Taking the initial conditions $\ln \sigma_{H,i,p+1}^2 \sim N(0, 10)$ we obtain a draw from the conditional distribution by using the Carter and Kohn (1994) simulation smoother.

Part b: V^T

The elements of the matrices $V_t = \text{diag}\left(\sigma_{V,1,t}^2, \ldots, \sigma_{V,M,t}^2\right)$, $t = p+1, \ldots, T$ are sampled in a similar manner as in part a. Conditional on the local mean process ψ^T we observe the innovations η_t for $t = p+2, \ldots, T$, since

$$\eta_t = \psi_t - \psi_{t-1}. \tag{28}$$

Given the initial conditions $\ln \sigma_{V,i,p+2}^2 \sim N(0, 10)$, we use the KSC algorithm to obtain the draws.

Part c: G^T

Let $t_{z,i}$ denote the index of the first available data point for the survey forecasts for variable i. Similarly as in parts a and b, conditional on ψ^T the innovations $g_{i,t}$, $t = t_{z,i}, \ldots, T$ are observable since

$$g_t = z_t - P_\psi \psi_t. \tag{29}$$

As above, we use the KSC algorithm to obtain the draws (with initial conditions $\ln \sigma_{G,i,t_{z,i}}^2 \sim N(0, 10)$).

2. Sampling from $p\left(\{\phi_H, \phi_V, \phi_G\} \mid \{\Lambda^T, V^T, G^T\}, A, B, \psi^T, S^T, \{Y, Z\}\right)$

A natural conjugate inverse Gamma prior is specified for the elements in the M-dimensional vector ϕ_H, with hyperparameters d_{ϕ_H} and $\underline{\phi}_H$. Then conditional on Λ^T we observe the innovations $u_{H,i,t}$, $i = 1, \ldots, M$, since

$$u_{H,i,t} = \ln \sigma_{H,i,t}^2 - \ln \sigma_{H,i,t-1}^2. \tag{30}$$

We can then sample the elements $\phi_{H,i}$, $i = 1, \ldots, M$, from

$$\phi_{H,i} \sim \text{IG}\left(\left(d_{\phi_H} \times \underline{\phi}_H + \sum_{t=p+2}^{T} u_{H,i,t}^2\right),\ (d_{\phi_H} + T - p - 1)\right). \tag{31}$$

Similarly, given the IG prior for ϕ_V with parameters d_{ϕ_V} and $\underline{\phi}_V$, $\phi_{V,i}$, $i = 1, \ldots, M$ is sampled from

$$\phi_{V,i} \sim \text{IG}\left(\left(d_{\phi_V} \times \underline{\phi}_V + \sum_{t=p+3}^{T} u_{V,i,t}^2\right),\ (d_{\phi_V} + T - p - 2)\right), \tag{32}$$

then for the M_z elements of ϕ_G we sample from

$$\phi_{G,i} \sim \text{IG}\left(\left(d_{\phi_G} \times \underline{\phi}_G + \sum_{t=t_{z,i}+1}^{T} u_{G,i,t}^2\right),\ (d_{\phi_G} + T - t_{z,i})\right). \tag{33}$$

3. Sampling from $p\left(A \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, B, \psi^T, S^T, \{Y, Z\}\right)$

Recall that conditional on B and the local mean process ψ^T the VAR innovations ε_t are observed and $\tilde{\varepsilon}_t = A\varepsilon_t$, where $\tilde{\varepsilon}_t \sim N(0, \Lambda_t)$. Recall also that the priors for the rows a_i, $i = 2, \ldots, M$ of A are Gaussian, i.e. $a_i \sim N\left(\underline{\mu}_{a,i}, \underline{\Omega}_{a,i}\right)$. As shown by Cogley and Sargent (2005), the posterior distributions for the rows a_i, $i = 2, \ldots, M$, of A are Gaussian:

$$a_i \sim N\left(\overline{\mu}_{a,i}, \overline{\Omega}_{a,i}\right). \tag{34}$$

For a detailed description of the derivation and the implementation of these conditional posterior distributions we refer to Cogley and Sargent (2005).

4. Sampling from $p\left(B \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, \psi^T, S^T, \{Y, Z\}\right)$

We use the recently proposed method of Carriero, Clark, and Marcellino (2016b) to sample the VAR coefficients B. Consider the Cholesky decomposition $\varepsilon_t = A^{-1}\Lambda^{0.5}\epsilon_t$, where ϵ_t is a vector of orthogonal shocks, i.e. $\epsilon_t \sim N(0, I)$. This decomposition can be written as:

$$\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{M,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ a_{2,1}^* & 1 & & \vdots \\ \vdots & & 1 & 0 \\ a_{M,1}^* & \ldots & a_{M,M-1}^* & 1 \end{bmatrix} \begin{bmatrix} \sigma_{H,1,t} & 0 & \ldots & 0 \\ 0 & \sigma_{H,2,t} & & \vdots \\ \vdots & & \ldots & 0 \\ 0 & \ldots & 0 & \sigma_{H,M,t} \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{M,t} \end{bmatrix}, \tag{35}$$

where $a_{i,j}^*$ denotes element (i, j) of A^{-1}. Next, we denote by $b^{(i,k)}$ the vector of VAR coefficients for equation i with lag k, i.e. the i^{th} row of B_k. Let j index the specific elements of the vector

$b^{(i,k)}$. Then we can rewrite the VAR equations as:

$$\begin{aligned}
y_{1,t} - \psi_{1,t} &= \sum_{k=1}^{p}\sum_{j=1}^{M} b_j^{(1,k)} \left(y_{j,t-k} - \psi_{j,t-k}\right) + \sigma_{H,1,t}\epsilon_{1,t}, \\
y_{2,t} - \psi_{2,t} &= \sum_{k=1}^{p}\sum_{j=1}^{M} b_j^{(2,k)} \left(y_{j,t-k} - \psi_{j,t-k}\right) + a^*_{2,1}\sigma_{H,1,t}\epsilon_{1,t} + \sigma_{H,2,t}\epsilon_{2,t}, \\
&\vdots \\
y_{M,t} - \psi_{M,t} &= \sum_{k=1}^{p}\sum_{j=1}^{M} b_j^{(M,k)} \left(y_{j,t-k} - \psi_{j,t-k}\right) + a^*_{M,1}\sigma_{H,1,t}\epsilon_{1,t} \\
&\quad + \ldots + a^*_{M,M-1}\sigma_{H,M-1,t}\epsilon_{M-1,t} + \sigma_{H,M,t}\epsilon_{M,t},
\end{aligned}$$

or in general for variable i:

$$\begin{aligned}
&\left(y_{i,t} - \psi_{i,t}\right) - \left(a^*_{i,1}\sigma_{H,1,t}\epsilon_{1,t} + \ldots + a^*_{i,i-1}\sigma_{H,i-1,t}\epsilon_{i-1,t}\right) \\
&\qquad = \sum_{k=1}^{p}\sum_{j=1}^{M} b_j^{(i,k)} \left(y_{j,t-k} - \psi_{j,t-k}\right) + \sigma_{H,i,t}\epsilon_{i,t}.
\end{aligned} \tag{36}$$

Given this representation the sampling from the conditional posterior distribution of B is performed by sampling iteratively the vectors $b^{(i,1:p)}$ for $i = 1, \ldots, M$. Note that Equation (36) is equivalent to a linear gaussian regression equation.[30] Then given the prior mean $\underline{\mu}_{B,i}$ and variance $\underline{\Omega}_{B,i}$ we use standard results for the posterior distributions in linear regression models to find the posterior mean $\overline{\mu}_{B,i}$ and the posterior variance $\overline{\Omega}_{B,i}$. After B is sampled we verify whether it satisfies the stability condition (by checking the eigenvalues of the companion form). If not step 4 is repeated.

We refer to Carriero, Clark, and Marcellino (2016b) for further derivations, which show that the resulting draw of B comes from the appropriate conditional posterior distribution.

5. Sampling from $p\left(\psi^T \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, S^T, \{Y, Z\}\right)$

We first show that we can cast the SLM model into a linear state space representation. We derive the representation for $p = 4$ (the lag length used in this paper), an adjustment for different values of p is straightforward. The measurement equation is given by:

$$\begin{bmatrix} z_t \\ y_t \end{bmatrix} = \begin{bmatrix} P_\psi & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \psi_t \\ y_t - \psi_t \\ y_{t-1} - \psi_{t-1} \\ y_{t-2} - \psi_{t-2} \\ y_{t-3} - \psi_{t-3} \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} \begin{bmatrix} g_t \end{bmatrix}, \tag{37}$$

[30] We condition on the first p time series observations $y_1, \ldots, y_p$. Furthermore in line with the conditioning on the initial observations we set the vectors $\psi_1, \ldots, \psi_p$ equal to the average over $y_1, \ldots, y_p$.

and the transition equation by:

$$\begin{bmatrix} \psi_t \\ y_t - \psi_t \\ y_{t-1} - \psi_{t-1} \\ y_{t-2} - \psi_{t-2} \\ y_{t-3} - \psi_{t-3} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & B_1 & B_2 & B_3 & B_4 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ y_{t-1} - \psi_{t-1} \\ y_{t-2} - \psi_{t-2} \\ y_{t-3} - \psi_{t-3} \\ y_{t-4} - \psi_{t-4} \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & I \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}, \quad (38)$$

where $g_t \sim N(0, G_t)$, $\eta_t \sim N(0, V_t)$, and $\varepsilon_t \sim N(0, H_t)$. Further g_t, η_t and ε_t are uncorrelated and we take a diffuse distribution for ψ_{p+1}. Conditional on the other parameters we have a linear Gaussian state space representation and we can take a draw from the conditional posterior distribution of ψ^T using a version of the simulation smoother of Carter and Kohn (1994). As the covariance matrix of the shocks in the transition equation is singular we adapt the simulation smoother following Kim and Nelson (1999, chap. 8.2).

We further deal with missing z_t values by setting the Kalman gain equal to zero when these are not available, for further implementation details we refer Durbin and Koopman (2012).

6. Sampling from $p\left(S^T \mid \{\Lambda^T, V^T, G^T\}, \{\phi_H, \phi_V, \phi_G\}, A, B, \psi^T, \{Y, Z\}\right)$

Conditional on the values of the other parameters $\tilde{\varepsilon}^2_{i,t}$, $\eta^2_{i,t}$ and $g^2_{i,t}$ are observable and we can sample the states, i.e. the elements of $S^T = \{s_H^T, s_V^T, s_G^T\}$, independently from the following discrete distributions:

$$p(s_{H,i,t} = j \mid \tilde{\varepsilon}^2_{i,t}, \sigma^2_{H,i,t}) \propto f_N(\ln(\tilde{\varepsilon}^2_{i,t} + 0.001) \mid \ln \sigma^2_{H,i,t} + m_j - 1.2704\ ,\ v_j^2), \quad (39)$$

$$p(s_{V,i,t} = j \mid \eta^2_{i,t}, \sigma^2_{V,i,t}) \propto f_N(\ln(\eta^2_{i,t} + 0.001) \mid \ln \sigma^2_{V,i,t} + m_j - 1.2704\ ,\ v_j^2), \quad (40)$$

$$p(s_{G,i,t} = j \mid g^2_{i,t}, \sigma^2_{G,i,t}) \propto f_N(\ln(g^2_{i,t} + 0.001) \mid \ln \sigma^2_{G,i,t} + m_j - 1.2704\ ,\ v_j^2), \quad (41)$$

where $j \in \{1, \ldots, 7\}$ indexes the states and $f_N(x \mid \mu, \sigma^2)$ denotes the normal probability density function with mean μ and variance σ^2 evaluated at the point x. For further details we refer to Kim, Shephard, and Chib (1998), who also provide the values for the state-dependent means, m_j, and variances, v_j^2.

B Data

Table A1: Description of the data

Variable	Transformation	Survey forecast	Small VAR	Source
euro area				
GDP (real)	$400 \times \Delta \ln$	1990Q1	+	Eurostat[1]
Consumption (private, real)	$400 \times \Delta \ln$	1990Q1	-	Eurostat[1]
Investment (total, real)	$400 \times \Delta \ln$	1998Q3	-	Eurostat[1]
Inflation (HICP)	$400 \times \Delta \ln$	1990Q1	+	ECB
Short-term interest rate	-	-	+	ECB[1]
United States				
GDP (real)	$400 \times \Delta \ln$	1990Q1	+	BEA
Consumption (personal, real)	$400 \times \Delta \ln$	1990Q1	-	BEA
Investment (business, real)	$400 \times \Delta \ln$	1998Q3	-	BEA
Inflation (CPI)	$400 \times \Delta \ln$	1990Q1	+	BLS
Short-term interest rate	-	-	+	FRB
Japan				
GDP (real)	$400 \times \Delta \ln$	1990Q1	+	CAO[2]
Consumption (private, real)	$400 \times \Delta \ln$	1990Q1	-	CAO[2]
Investment (business, real)	$400 \times \Delta \ln$	1998Q3	-	OECD
Inflation (CPI)	$400 \times \Delta \ln$	1990Q1	+	MIC[2]
Short-term interest rate	-	-	+	OECD

Notes: Description of the data used in the forecast evaluations. The third column indicates the first quarter for which the long-term survey forecast is available. The fourth column indicates whether the variable is included in the three-variable VAR.
[1]: Official data backdated using the Area Wide Model database (Fagan, Henry, and Mestre, 2001); [2]: Official data backdated using the OECD database.

C Additional results

Table A2: Forecast evaluation results for five-variable VARs

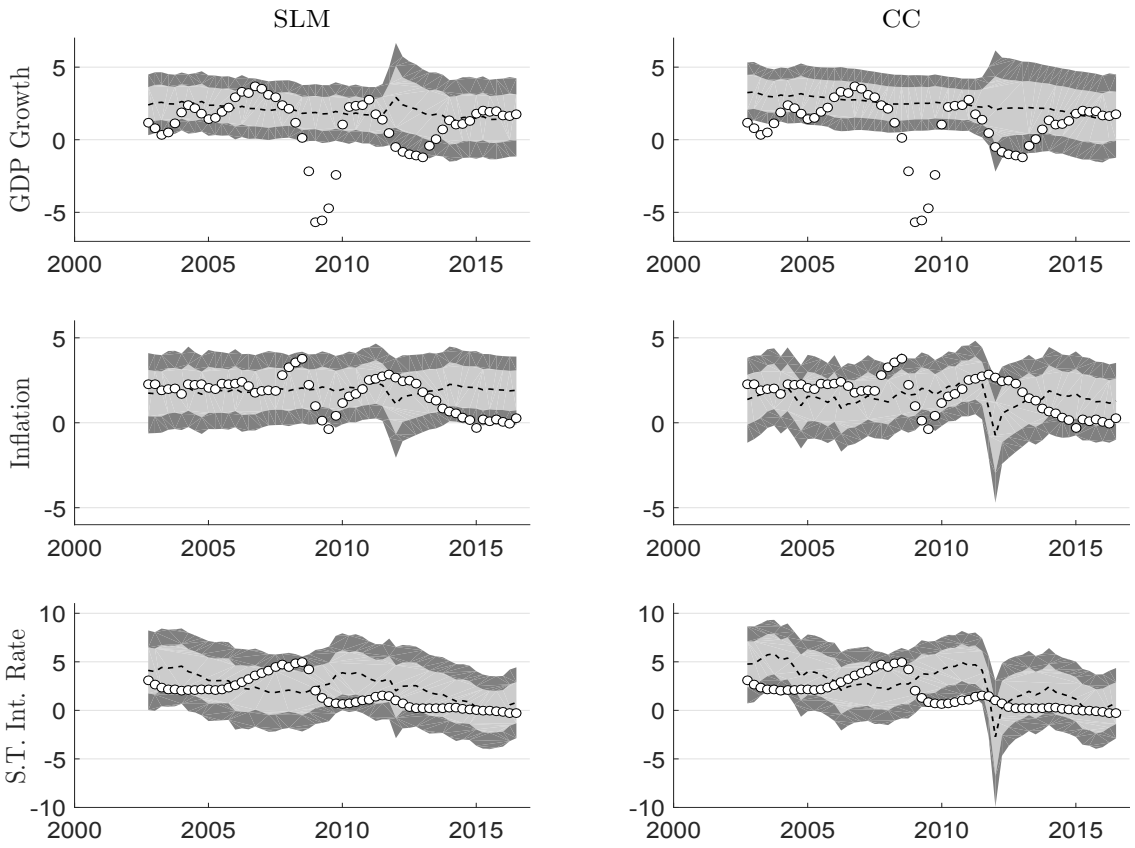
RMSE		euro area CC	euro area DP	euro area SLM	United States CC	United States DP	United States SLM	Japan CC	Japan DP	Japan SLM
GDP	1Q	2.23	0.95**	**0.95****	2.32	**0.98***	0.99	4.33	**1.00**	1.02°°
Growth	4Q	2.14	0.93**	**0.88***	1.90	**0.90****	0.95*	2.62	**0.98**	1.00
	8Q	2.48	0.92**	**0.87***	2.19	**0.91***	0.92**	2.56	**1.00**	1.00
	12Q	2.51	0.92***	**0.92***	2.21	**0.93****	0.95***	2.50	**0.98**	1.00
Inflation	1Q	1.29	1.00	**0.99**	2.27	**0.99**	1.00	1.50	1.00	**0.97**
	4Q	1.03	0.97	**0.96**	1.27	1.03°°	**1.00**	**1.05**	1.01	1.02
	8Q	1.27	0.98	**0.91**	1.37	1.04°°	**0.93**	**1.15**	1.03	1.05
	12Q	1.29	0.94	**0.88**	1.41	1.01	**0.93**	**1.15**	1.01	1.03
S.T. Int.	1Q	0.30	3.37°°°	**0.94**	0.42	2.39°°°	**0.98**	**0.13**	4.20°°°	1.28°°°
Rate	4Q	1.08	1.39°°°	**0.82**	**1.33**	1.13	1.01	**0.34**	1.80°°°	1.12
	8Q	1.87	1.13°°	**0.78**	2.26	**0.91**	0.97	**0.48**	1.40°°	1.26°°°
	12Q	2.24	1.03	**0.79***	2.71	**0.88**	0.89	**0.54**	1.46°°°	1.46°°°
LPS		**CC**	**DP**	**SLM**	**CC**	**DP**	**SLM**	**CC**	**DP**	**SLM**
GDP	1Q	2.17	−0.04***	**−0.05*****	2.24	**−0.03*****	−0.01	2.88	**0.00**	0.02°°
Growth	4Q	2.85	−0.16**	**−0.23****	2.12	**−0.15*****	−0.07	2.57	**−0.05**	0.03
	8Q	3.39	−0.34**	**−0.57***	2.40	**−0.18****	−0.16**	2.58	**−0.01**	0.04
	12Q	3.23	−0.27***	**−0.47****	2.40	**−0.11****	−0.11**	2.51	**−0.03*****	0.03
Inflation	1Q	1.73	−0.01	**−0.04**	2.35	**−0.02****	0.00	1.84	−0.03	**−0.03**
	4Q	1.46	−0.03	**−0.05**	1.68	**−0.02**	0.04	1.53	−0.02	**−0.02**
	8Q	1.62	**−0.05**	−0.05	1.83	−0.04	**−0.08**	**1.62**	0.00	0.02
	12Q	1.66	**−0.06**	−0.05	1.88	−0.04**	**−0.11**	1.59	**−0.04**	0.04
S.T. Int.	1Q	0.33	**−0.01***	0.39°°°	0.59	**−0.03*****	0.10°°°	−0.09	**−0.01****	0.27°°°
Rate	4Q	1.47	**−0.03*****	0.08	1.75	**−0.12***	−0.03	0.78	**0.00**	0.09°°°
	8Q	2.04	−0.05***	**−0.13**	2.48	**−0.30***	−0.23	1.22	**0.00**	0.05°°°
	12Q	2.26	−0.08***	**−0.20****	2.77	**−0.44****	−0.42**	1.44	**0.00**	0.05°°°

Notes: The VAR includes GDP growth, consumption growth, investment growth, inflation and the short-term interest rate. See notes to Table 1 for further information.

Figure A1: 12-quarter-ahead density forecasts for inflation

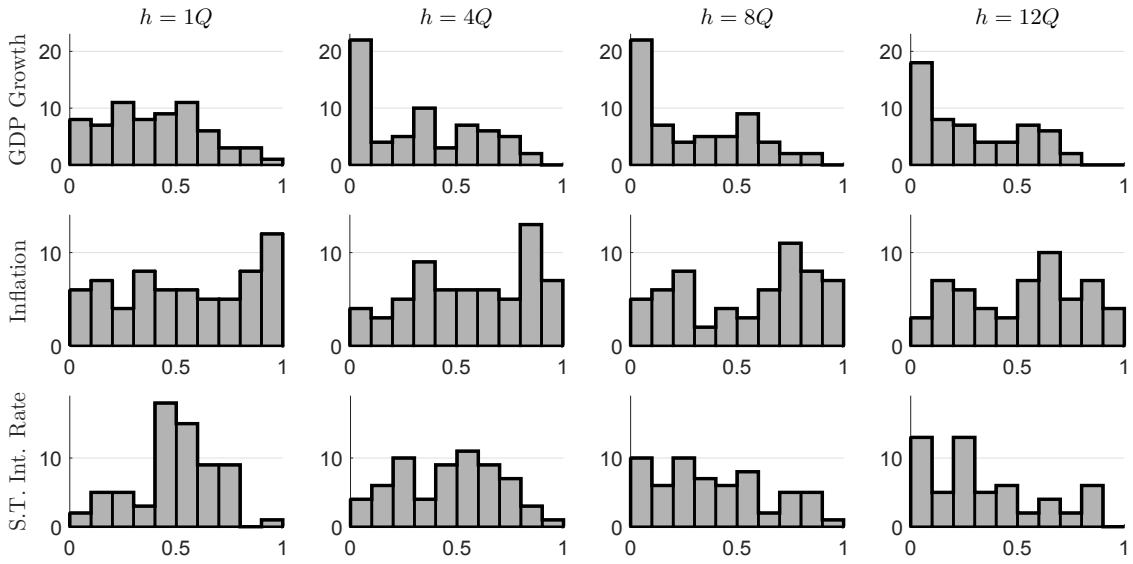
Notes: Predictive distributions for euro area inflation 12 quarters ahead for the SLM model with three variables (left panel) and the (univariate) UCSV model (right panel). Average rates over the third year are shown (both for the forecasts and the realisations). The x-axes indicate the time of the realisations. The predictive distributions are summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. Realisations are indicated by white circles.

Figure A2: 12-quarter-ahead density forecasts



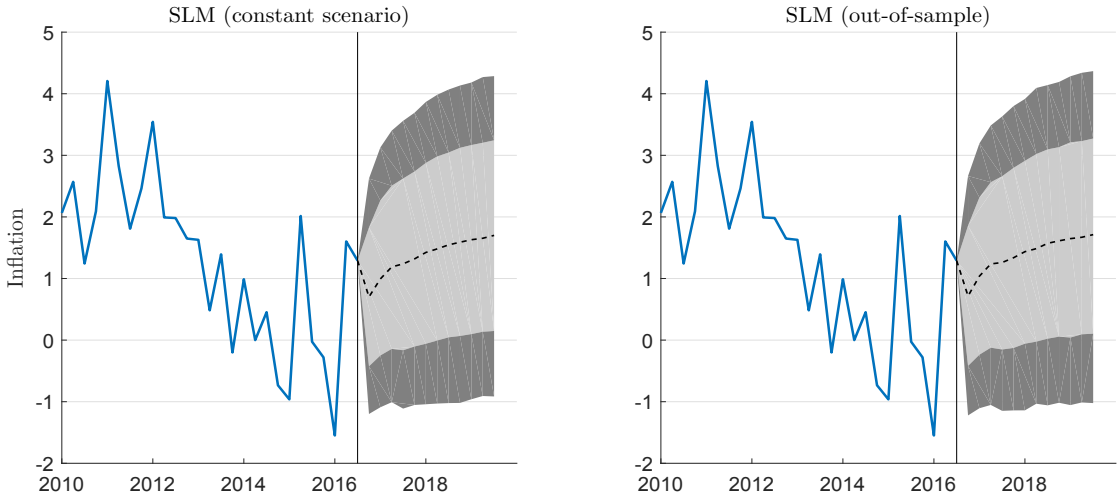
Notes: Predictive distributions for the SLM (left panel) and CC (right panel) models with three variables for the euro area for a forecast horizon of 12 quarters. Average rates over the third year are shown (both for the forecasts and the realisations). The x-axes indicate the time of the realisations. The predictive distributions are summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. Realisations are indicated by white circles.

Figure A3: Probability integral transforms for the CC model



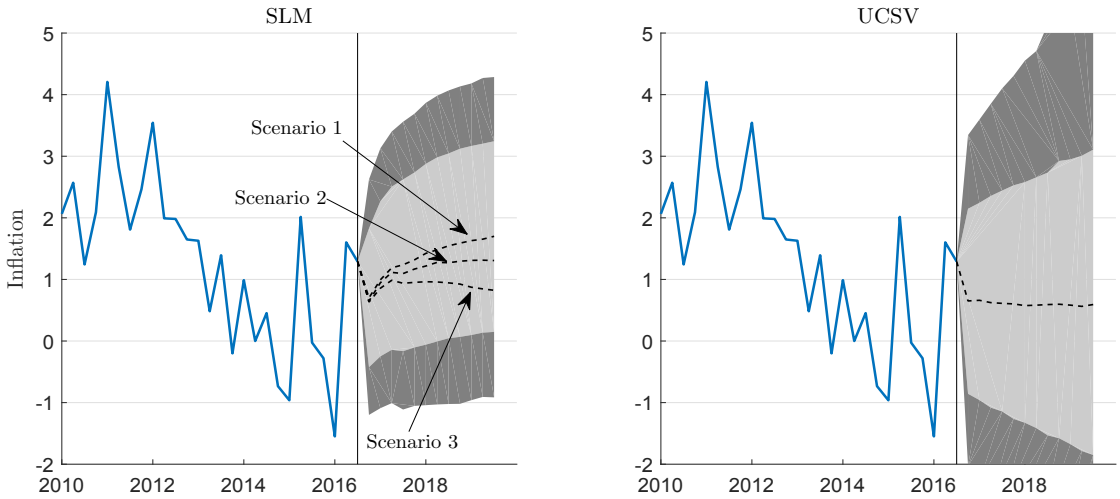
Notes: The histograms show the PITs for the CC model with three variables for the euro area. The columns correspond to the forecast horizons and the rows to the variables.

Figure A4: Conditional and unconditional forecasts from the SLM model for inflation



Notes: 12-quarter-ahead forecasts for euro area inflation for the SLM model with three variables. The in-sample period ends at 2016Q3 and the forecast period is 2016Q4 to 2019Q3. The left panel shows the forecasts conditional on the scenario that keeps the long-term expectations constant out-of-sample. The right panel shows the unconditional forecasts. The forecast distribution is summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands.

Figure A5: (Conditional) forecasts for inflation



Notes: 12-quarter-ahead forecasts for euro area inflation for the SLM model with three variables and the UCSV model. The in-sample period ends at 2016Q3 and the forecast period is 2016Q4 to 2019Q3. The forecasts for the SLM model are conditional on different scenarios for future developments in the long-term survey forecasts: 1 - kept at the value of 2016Q3; 2 - linearly declining to 1.5% in 2019Q3; 3 - linearly declining to 1% in 2019Q3. For the UCSV model the forecasts are unconditional. The forecast distribution is summarised by the median (dashed line) and the 68% (light grey area) and 90% (dark grey area) coverage bands. For the SLM model, the distribution is shown for scenario 1, for the remaining scenarios only the median is reported.

Table A3: Forecast evaluation results for three-variable VARs, rolling window estimation

		euro area			United States			Japan		
RMSE		CC	DP	SLM	CCR	DP	SLM	CC	DP	SLM
GDP	1Q	2.17	**0.93*****	0.95**	2.66	0.98	**0.96****	**4.46**	1.00	1.01
Growth	4Q	2.12	0.91**	**0.90**	2.01	0.93**	**0.89***	2.53	**1.00**	1.01
	8Q	2.46	0.92***	**0.90**	2.13	0.95*	**0.91***	**2.50**	1.03$^{\circ\circ}$	1.02
	12Q	2.53	**0.93*****	0.93**	2.16	0.98*	**0.94***	2.48	**0.98***	1.00
Inflation	1Q	1.30	1.00	**0.98**	2.29	**1.00**	1.00	1.57	**1.00**	1.02
	4Q	1.05	0.98	**0.94****	**1.29**	1.01	1.01	**1.04**	1.03	1.11
	8Q	1.23	0.94**	**0.93**	1.38	0.99	**0.91***	**1.15**	1.10	1.06
	12Q	1.23	0.97	**0.92**	1.45	1.01	**0.91*****	1.16	0.98	**0.95**
S.T. Int.	1Q	**0.31**	3.44$^{\circ\circ\circ}$	1.10$^{\circ\circ}$	**0.37**	2.41$^{\circ\circ\circ}$	1.08°	**0.12**	7.89$^{\circ\circ\circ}$	1.78°
Rate	4Q	1.12	1.24°	**0.99**	**1.30**	1.11$^{\circ\circ}$	1.02	**0.29**	2.93$^{\circ\circ\circ}$	1.25$^{\circ\circ}$
	8Q	1.85	1.04	**0.93**	2.32	**0.93**	0.94	**0.40**	2.40$^{\circ\circ\circ}$	1.42$^{\circ\circ}$
	12Q	2.21	1.06	**0.89****	2.86	0.91*	**0.87****	**0.43**	2.53$^{\circ\circ\circ}$	1.65$^{\circ\circ}$
LPS		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.14	−0.03***	**−0.06****	2.36	−0.04***	**−0.06****	**2.94**	0.00	0.02°
Growth	4Q	2.84	−0.14**	**−0.37**	2.27	−0.16**	**−0.21***	2.66	0.03	**−0.01**
	8Q	3.33	−0.25**	**−0.46***	2.60	−0.14***	**−0.29***	**2.67**	0.04	0.02
	12Q	3.15	−0.19***	**−0.33*****	2.65	−0.07***	**−0.26**	2.49	**−0.01**	0.05
Inflation	1Q	1.75	−0.01	**−0.04**	2.43	**−0.02**	0.01	1.93	**−0.04**	−0.02
	4Q	1.67	−0.01	**−0.09**	1.77	**−0.01**	0.05	1.57	**−0.07**	−0.01
	8Q	1.96	−0.01	**−0.05**	1.90	−0.05	**−0.09***	1.71	−0.04	**−0.07**
	12Q	1.89	0.01	**−0.02**	2.00	−0.07	**−0.10***	1.67	**−0.14**	−0.09
S.T. Int.	1Q	0.21	**−0.01***	0.20$^{\circ\circ\circ}$	0.37	**−0.03****	0.13$^{\circ\circ\circ}$	**−0.42**	0.01$^{\circ\circ}$	0.32$^{\circ\circ\circ}$
Rate	4Q	1.49	**−0.01**	−0.01	1.77	**−0.12***	−0.04	**0.39**	0.09$^{\circ\circ}$	0.22$^{\circ\circ\circ}$
	8Q	2.07	−0.02*	**−0.06**	2.64	**−0.29***	−0.26	**0.79**	0.17$^{\circ\circ}$	0.27$^{\circ\circ}$
	12Q	2.25	−0.04**	**−0.12****	3.00	−0.43**	**−0.51****	**1.01**	0.18	0.33$^{\circ\circ}$

Notes: Forecast evaluation results for the implementation where parameters are estimated over a rolling window of 80 quarters. See notes to Table 1 for further information.

Table A4: Forecast evaluation results for three-variable VARs, constant parameters out-of-sample

		euro area			United States			Japan		
RMSE		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.26	0.96**	**0.94*****	2.62	**0.96***	0.98	**4.43**	1.01	1.01
Growth	4Q	2.16	0.95**	**0.86***	2.03	**0.88****	0.92*	2.57	**1.00**	1.00
	8Q	2.47	0.93***	**0.87****	2.21	0.92**	**0.89***	**2.52**	1.02	1.01
	12Q	2.51	0.92***	**0.90*****	2.24	0.93**	**0.93****	2.48	**0.98**	1.01°
Inflation	1Q	1.27	**0.98**	0.99	2.25	1.00	**0.99**	1.56	1.00	**0.99**
	4Q	1.01	0.97	**0.95**	1.26	0.99	**0.98**	**1.06**	1.06	1.03
	8Q	1.21	0.97	**0.92**	1.37	0.97	**0.93**	**1.15**	1.03	1.05
	12Q	1.22	0.96	**0.92**	1.41	0.99	**0.93**	**1.16**	1.03	1.02
S.T. Int.	1Q	**0.28**	3.63$^{\circ\circ\circ}$	1.00	0.41	2.19$^{\circ\circ\circ}$	**0.97**	**0.12**	5.55$^{\circ\circ\circ}$	1.32$^{\circ\circ\circ}$
Rate	4Q	1.05	1.40$^{\circ\circ\circ}$	**0.88**	1.36	1.04	**1.00**	**0.31**	2.62$^{\circ\circ\circ}$	1.18$^{\circ\circ}$
	8Q	1.80	1.20$^{\circ\circ\circ}$	**0.83**	2.34	**0.90**	0.95	**0.43**	2.28$^{\circ\circ\circ}$	1.34$^{\circ\circ}$
	12Q	2.17	1.17$^{\circ\circ\circ}$	**0.83***	2.81	**0.87**	0.88*	**0.49**	2.26$^{\circ\circ\circ}$	1.53$^{\circ\circ\circ}$
LPS		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.18	−0.04***	**−0.07*****	2.36	**−0.04*****	−0.03	2.90	**−0.01**	0.02
Growth	4Q	2.90	−0.17**	**−0.32****	2.21	**−0.17*****	−0.10	2.54	**−0.03**	0.07
	8Q	3.44	−0.32**	**−0.58***	2.43	−0.17***	**−0.20*****	2.58	**0.00**	0.11
	12Q	3.30	−0.22***	**−0.45****	2.44	−0.10***	**−0.15*****	2.52	**−0.01****	0.11
Inflation	1Q	1.73	0.00	**−0.03**	2.35	**−0.02**	−0.02	1.90	−0.01	**−0.01**
	4Q	1.45	−0.01	**−0.06**	1.68	**−0.01**	0.01	1.55	**−0.01**	−0.01
	8Q	1.59	−0.03	**−0.04**	1.83	−0.05	**−0.06**	1.64	**−0.01**	0.02
	12Q	1.61	**−0.04**	−0.03	1.90	**−0.06**	−0.06	1.61	**−0.04**	0.01
S.T. Int.	1Q	0.30	**−0.01***	0.39$^{\circ\circ\circ}$	0.58	**−0.03*****	−0.01	−0.15	**0.00**	0.17$^{\circ\circ\circ}$
Rate	4Q	1.46	**−0.02*****	0.05	1.79	**−0.14****	−0.04	0.72	0.01	**−0.02**
	8Q	2.04	−0.04***	**−0.17**	2.60	**−0.32****	−0.07	1.15	0.03	**−0.04***
	12Q	2.25	−0.05***	**−0.23*****	2.92	**−0.46****	−0.20	1.37	0.03	**−0.02**

Notes: Forecast evaluation results for the implementation in which the time-varying parameters are kept constant out-of-sample (at the values from the last point of the estimation sample). See notes to Table 1 for further information.

Table A5: Forecast evaluation results for three-variable VARs, constant variance

		euro area			United States			Japan		
RMSE		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.30	0.98^{*}	$\mathbf{0.94}^{***}$	2.63	$\mathbf{0.97}^{**}$	0.98	4.46	**0.99**	1.00
Growth	4Q	2.19	0.95^{*}	$\mathbf{0.85}^{*}$	2.02	$\mathbf{0.90}^{**}$	0.93^{*}	2.60	1.00	**1.00**
	8Q	2.47	0.91^{**}	$\mathbf{0.86}^{**}$	2.20	0.91^{*}	$\mathbf{0.90}^{*}$	**2.52**	1.01	1.01
	12Q	2.50	0.91^{***}	$\mathbf{0.90}^{***}$	2.22	$\mathbf{0.93}^{**}$	0.93^{**}	2.49	**0.99**	1.00
Inflation	1Q	1.27	1.03	**0.99**	2.25	**0.99**	0.99	1.56	1.03	**0.99**
	4Q	1.02	1.03	**0.96**	1.26	**0.95**	0.98	**1.06**	1.08	1.04
	8Q	1.21	1.00	**0.93**	1.38	0.99	**0.92**	**1.15**	1.07	1.05
	12Q	1.21	1.00	**0.93**	1.41	0.96^{*}	**0.93**	**1.16**	1.08	1.03
S.T. Int.	1Q	**0.28**	$4.36^{\circ\circ\circ}$	1.02	0.44	$1.88^{\circ\circ\circ}$	**0.94**	**0.13**	$5.46^{\circ\circ\circ}$	$1.35^{\circ\circ\circ}$
Rate	4Q	1.05	$1.42^{\circ\circ\circ}$	**0.88**	1.39	0.99	**0.99**	**0.33**	$2.56^{\circ\circ\circ}$	1.17°
	8Q	1.82	$1.11^{\circ\circ}$	**0.81**	2.35	$\mathbf{0.89}^{*}$	0.95	**0.46**	$2.12^{\circ\circ\circ}$	$1.33^{\circ\circ\circ}$
	12Q	2.21	1.03	$\mathbf{0.80}^{*}$	2.80	0.88^{*}	$\mathbf{0.87}^{*}$	**0.52**	$2.03^{\circ\circ\circ}$	$1.53^{\circ\circ\circ}$
LPS		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.23	-0.04^{***}	$\mathbf{-0.08}^{***}$	2.37	$\mathbf{-0.04}^{***}$	-0.03	2.93	$\mathbf{-0.01}$	0.02
Growth	4Q	2.83	-0.19^{**}	$\mathbf{-0.33}^{**}$	2.22	$\mathbf{-0.17}^{***}$	-0.10	2.50	$\mathbf{-0.03}$	0.05
	8Q	3.26	-0.30^{**}	$\mathbf{-0.55}^{**}$	2.40	-0.16^{***}	$\mathbf{-0.21}^{**}$	2.50	**0.00**	0.07
	12Q	3.17	-0.23^{***}	$\mathbf{-0.47}^{**}$	2.38	-0.09^{***}	$\mathbf{-0.15}^{**}$	2.44	$\mathbf{-0.02}^{**}$	0.04
Inflation	1Q	1.74	0.00	**−0.04**	2.41	0.00	**−0.01**	1.89	**−0.01**	−0.01
	4Q	1.47	−0.01	**−0.06**	1.70	**−0.01**	0.00	1.53	0.01	**0.00**
	8Q	1.60	**−0.03**	−0.02	1.84	−0.06	**−0.10**	**1.62**	0.00	0.01
	12Q	1.60	**−0.03**	0.01	1.88	-0.06^{*}	**−0.09**	1.60	**−0.03**	0.02
S.T. Int.	1Q	0.35	$\mathbf{-0.01}^{*}$	$0.52^{\circ\circ\circ}$	0.68	$\mathbf{-0.03}^{***}$	$0.08^{\circ\circ}$	**−0.02**	0.00	$0.26^{\circ\circ\circ}$
Rate	4Q	1.49	$\mathbf{-0.03}^{***}$	0.16	1.79	$\mathbf{-0.14}^{**}$	-0.07	**0.84**	0.01	$0.07^{\circ\circ\circ}$
	8Q	2.04	-0.05^{***}	**−0.09**	2.53	$\mathbf{-0.30}^{*}$	-0.28	**1.26**	0.02	$0.03^{\circ\circ}$
	12Q	2.25	-0.06^{***}	$\mathbf{-0.16}^{*}$	2.82	-0.43^{**}	$\mathbf{-0.47}^{**}$	**1.46**	0.02	$0.04^{\circ\circ}$

Notes: Forecast evaluation results for model versions with constant variance. See notes to Table 1 for further information.

Table A6: Forecast evaluation results for three-variable VARs, robust evaluation criteria

		euro area			United States			Japan		
MAE		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	1.43	0.97*	**0.95***	1.98	**0.97**	0.99	**3.04**	1.01	1.02°
Growth	4Q	1.46	0.93*	**0.84****	1.55	0.89**	**0.89****	1.72	**0.99**	1.01
	8Q	1.67	0.93	**0.85****	1.47	0.86***	**0.84****	**1.62**	1.03	1.03
	12Q	1.67	0.93***	**0.89****	1.48	0.89***	**0.85*****	**1.56**	1.05$^{\circ\circ}$	1.02
Inflation	1Q	0.96	1.03	**0.99**	**1.45**	1.02	1.01	**1.06**	1.03	1.04
	4Q	0.76	1.03	**0.97**	**0.95**	1.03	1.01	**0.72**	1.21$^{\circ\circ}$	1.20$^{\circ\circ}$
	8Q	1.02	0.98	**0.90**	1.05	0.97	**0.95**	**0.78**	1.12	1.11
	12Q	0.99	0.98	**0.88**	1.11	**0.92***	0.93	**0.80**	1.11	1.14°
S.T. Int.	1Q	**0.20**	5.23$^{\circ\circ\circ}$	1.11°	**0.29**	2.16$^{\circ\circ\circ}$	1.01	**0.08**	6.98$^{\circ\circ\circ}$	1.52$^{\circ\circ\circ}$
Rate	4Q	0.73	1.61$^{\circ\circ\circ}$	**0.99**	1.09	**0.99**	1.01	**0.23**	2.74$^{\circ\circ\circ}$	1.31$^{\circ\circ}$
	8Q	1.42	1.21$^{\circ\circ\circ}$	**0.87**	2.01	**0.86**	0.94	**0.35**	1.95$^{\circ\circ\circ}$	1.47$^{\circ\circ}$
	12Q	1.87	1.08°	**0.83****	2.52	**0.82***	0.85*	**0.40**	2.00$^{\circ\circ\circ}$	1.69$^{\circ\circ}$
LPS (t)		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	2.11	−0.04***	**−0.06*****	2.35	**−0.04*****	−0.03	**2.84**	0.00	0.01
Growth	4Q	2.45	−0.14**	**−0.23***	2.19	**−0.16*****	−0.12*	2.38	**−0.01**	0.04
	8Q	2.74	−0.22***	**−0.35****	2.28	−0.16***	**−0.18*****	2.39	**0.00**	0.06
	12Q	2.62	−0.17***	**−0.25*****	2.28	−0.11***	**−0.13*****	2.35	**−0.03****	0.01
Inflation	1Q	1.68	0.00	**−0.02**	2.13	**−0.02**	0.00	**1.83**	0.01	0.01
	4Q	1.42	−0.02	**−0.05**	1.64	**0.00**	0.01	**1.51**	0.03	0.01
	8Q	1.60	−0.03	**−0.04**	1.77	−0.03	**−0.06**	**1.59**	0.02	0.04
	12Q	1.61	**−0.03**	**−0.03**	1.83	−0.04	**−0.08**	1.56	**0.00**	0.05
S.T. Int.	1Q	0.27	**−0.01***	0.46$^{\circ\circ\circ}$	0.55	**−0.04*****	0.08$^{\circ\circ\circ}$	−0.21	**0.00**	0.32$^{\circ\circ\circ}$
Rate	4Q	1.45	**−0.02*****	0.11	1.77	**−0.14****	−0.05	**0.67**	0.02	0.14$^{\circ\circ\circ}$
	8Q	2.04	−0.04**	**−0.12**	2.56	**−0.29****	−0.25	**1.10**	0.04°	0.10$^{\circ\circ\circ}$
	12Q	2.26	−0.05***	**−0.18****	2.87	−0.41**	**−0.44****	**1.31**	0.06	0.11$^{\circ\circ\circ}$
CRPS		CC	DP	SLM	CC	DP	SLM	CC	DP	SLM
GDP	1Q	1.10	−0.05***	**−0.07****	1.42	**−0.05*****	−0.04	2.26	**0.00**	0.02
Growth	4Q	1.13	−0.09**	**−0.17****	1.13	**−0.14*****	−0.12**	1.30	**0.00**	0.01
	8Q	1.33	−0.14***	**−0.22****	1.16	−0.14***	**−0.16****	**1.28**	0.01	0.03
	12Q	1.31	−0.13***	**−0.15*****	1.16	−0.10***	**−0.12*****	1.25	**−0.02***	0.00
Inflation	1Q	0.70	0.00	**−0.01**	1.08	**0.00**	0.01	**0.80**	0.02	0.02
	4Q	0.55	−0.01	**−0.02**	**0.69**	0.00	0.00	**0.57**	0.05	0.04
	8Q	0.68	−0.02	**−0.04**	0.76	−0.02	**−0.05**	**0.62**	0.03	0.04
	12Q	0.68	−0.02	**−0.04**	0.79	−0.02	**−0.05**	**0.62**	0.02	0.04
S.T. Int.	1Q	0.16	**0.00**	0.06$^{\circ\circ\circ}$	0.22	**−0.01*****	0.01	0.09	**0.00**	0.03$^{\circ\circ\circ}$
Rate	4Q	0.55	**−0.01***	0.02	0.76	**−0.08****	0.00	**0.21**	0.01	0.04$^{\circ\circ\circ}$
	8Q	1.02	−0.03*	**−0.14**	1.43	**−0.23***	−0.12	**0.32**	0.04$^{\circ\circ}$	0.06$^{\circ\circ\circ}$
	12Q	1.28	−0.06**	**−0.23****	1.79	**−0.35****	−0.31*	**0.40**	0.07$^{\circ\circ}$	0.09$^{\circ\circ\circ}$

Notes: Forecast evaluation results based on robust criteria. The top panel displays the mean absolute error (MAE), the middle panel the LPS based on a t-distribution and the bottom panel the continuous ranked probability score (CRPS). As in Table 1 the accuracy measures are reported in absolute terms for the CC model and relative to the CC model for the remaining two models (for the CRPS we subtract the CRPS value of the CC model from those of the DP and SLM models). See notes to Table 1 for further information.

Table A7: Forecast evaluation results for the SLM model with three and 15-variables, comparison to survey forecasts

		euro area			United States			Japan		
RMSE		Survey	SLM 3	SLM 15	Survey	SLM 3	SLM 15	Survey	SLM 3	SLM 15
GDP	1Q	2.55	**0.83***	0.85*	2.62	0.98	**0.86***	**4.31**	1.03°	1.01
Growth	4Q	2.09	**0.89***	0.90*	1.93	0.97	**0.94**	**2.52**	1.02	1.05°
	8Q	2.15	0.99	**0.99**	2.00	**0.98**	1.00	2.61	**0.98**	1.00
	12Q	**2.22**	1.02	1.01	2.08	**0.99****	1.01	2.55	**0.98**	0.99*
Inflation	1Q	1.39	**0.90***	0.92	2.25	0.99	**0.95**	1.94	**0.80*****	0.86***
	4Q	1.01	**0.95**	0.96	1.27	0.98	**0.97**	1.54	**0.71****	0.89*
	8Q	**1.06**	1.06	1.05	1.27	**0.99**	1.00	1.54	**0.79***	0.96
	12Q	**1.09**	1.03	1.02	1.32	**1.00**	1.00	1.47	**0.80**	0.97***

Notes: Point forecast evaluation results for long-term survey forecasts and the SLM models with three and 15 variables. The RMSEs are reported in absolute terms for the survey forecasts and relative to the survey forecasts for the SLM models. See notes to Table 1 for further information.